Transactions of the ASME

TRANSACTIONS OF THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS

VOLUME 79 MAY 1957 NUMBER 4

Transactions

of The American Society of Mechanical Engineers

Published on the tenth of every month, except March, June, September, and December

Published monthly by The American Society of Mechanical Engineers. Publication office at 20th and Northampton Streets, Easton, Pa. The editorial department is located at the headquarters of the Society, 29 West Thirty-Ninth Street, New York 18, N. Y. Cable address, "Mechaneer," New York. Price $1.50 a copy, $12.00 annually for Transactions and the *Journal of Applied Mechanics;* to members, $1.00 a copy, $6.00 annually. Add $1.50 for postage to all countries outside the United States, Canada, and Pan American Union. Changes of address must be received at Society headquarters seven weeks before they are to be effective on the mailing list. Please send old as well as new address.... By-Law: The Society shall not be responsible for statements or opinions advanced in papers or ... printed in its publications (B13, Par. 4).... Entered as second-class matter March 2, 1928, at the Post Office at Easton, Pa., under the Act of August 24, 1912....

Strength Characteristics of Rock Samples Under Hydrostatic Pressure

By R. O. BREDTHAUER,[1] HOUSTON, TEXAS

Other investigations have shown that rock cylinders exposed to hydrostatic pressures of 10,000 atm and loaded axially in compression behave in a ductile fashion as contrasted with their brittle behavior when tested in axial compression at atmospheric pressure. Similar tests were made by the author on rock cylinders in the pressure range 0 to 15,000 psi, which is equivalent to pressures exerted by drilling muds upon the face of formations exposed to the cutting action of rock-bit teeth. Results indicated that many formations commonly encountered in drilling operations undergo a brittle-to-ductile behavior transition in the range 0 to 15,000-psi hydrostatic pressure. The great majority of tests were made on samples which were dry and which were protected from the surrounding fluid with a plastic jacket. At atmospheric pressure, all formations so tested were brittle. The compressive strength of all formations tested in this condition increased with increasing confining pressure.

Introduction

FOR many years, it has been known that laboratory drilling rate tests did not agree with results of field tests, even when it was known that the formations were the same. Although a number of variables could have contributed to this result, it was generally felt that fluid pressure exerted by the mud column above the formation being drilled must have an important effect. Thus, a bit designed to perform quite well in the laboratory under atmospheric pressure may perform unsatisfactorily and uneconomically in the field under more practical conditions.

The stress imposed by fluid-column pressure is not the only stress to which a formation at the bottom of an oil-well bore hole is exposed. The original stress state of the formation before the hole is drilled is probably compressive in the three principal directions, with the principal stresses being unequal. Introduction of a hole relieves the compressive stress at the wall of the hole. At the intersection of the wall and the bottom, the stress condition becomes much more complex. Addition of a fluid column into the hole superimposes other compressive stresses upon those already existing. Finally, action of the rock-bit teeth subjects the formation to additional stresses.

Even though the stress condition at the bottom of the hole is extremely complex, the principal stresses are probably all compressive. Before the rock-bit teeth bear upon the formation surface, a rough approximation of the stress state may be uniform hydrostatic stress. It is believed that a study of rock failure under conditions of uniform compressive stress in two principal directions and a higher compressive stress in the third principal direction will aid in an understanding of rock-failure mechanics in oil-well drilling.

Previous experimental investigations (1–7)[2] have shown that cylindrical rock samples undergo marked changes in strength characteristics when tested under high hydrostatic confining pressures as compared to their behavior when tested under atmospheric pressure. Probably the first work in which rock cylinders were exposed to hydrostatic confining pressure while being loaded axially in compression was that of von Karman (1) in 1911. Generally, the cylindrical rock samples support a much greater axial load when subjected to high confining pressures than when subjected to atmospheric pressure only. Also, some formations which failed in a brittle manner at atmospheric pressure failed in a ductile fashion when tested under high confining pressures. Much of the previous work was aimed at explanation of rock displacement caused by geologic phenomena, and accordingly tests were made at high pressures (10,000 atm) far exceeding the pressure range which is of interest in the drilling of oil-well holes.

The purpose of this work is to study the effect of hydrostatic pressures in the range 0 to 15,000 psi upon cylindrical rock samples when loaded in axial compression.

Definitions

Confining pressure. Hydrostatic pressure exerted on all surfaces of the cylindrical specimen by the fluid in the testing machine.

Differential stress. Difference between the hydrostatic stress and the axial compressive stress. Calculation of the differential stress is based upon the assumption that volume of the specimen remains constant during the test, and that lateral deformation is uniform over the entire length of the specimen.

Flow. That deformation which is not instantly recoverable, except fracture.

Fracture. Deformation with loss of cohesion.

Yield strength. Differential stress at which a specimen begins to flow, or the differential stress at which the stress-strain curve shows deviation from purely elastic behavior.

Strength. Maximum differential compressive stress required to fracture specimens which did not flow, and the maximum differential stress imposed on those specimens which did flow.

Brittle material or brittle formation. Characterized by fracture while being deformed elastically. In speaking of a brittle material, the test condition must also be specified, as a material may be brittle under one set of conditions and ductile under another set.

Strain. Axial deformation experienced by a specimen, expressed in per cent of its original length.

Formation names. Generally, terms in common use to indicate a locality, such as "Carthage" marble, "Knippa" basalt.

Equipment

The pressure-test equipment (Fig. 1) used was based upon a de-

[1] Project Engineer, Metallurgical Department, Hughes Tool Company.

Contributed by the Petroleum Division and presented at the Petroleum-Mechanical Engineering Conference, Dallas, Texas, September 23–26, 1956, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, July 19, 1956. Paper No. 56—PET-23

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

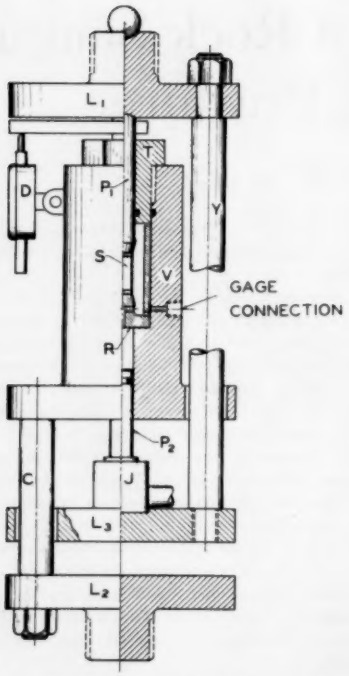


FIG. 1(*a*) SCHEMATIC DIAGRAM OF HIGH-PRESSURE STRENGTH-TEST APPARATUS

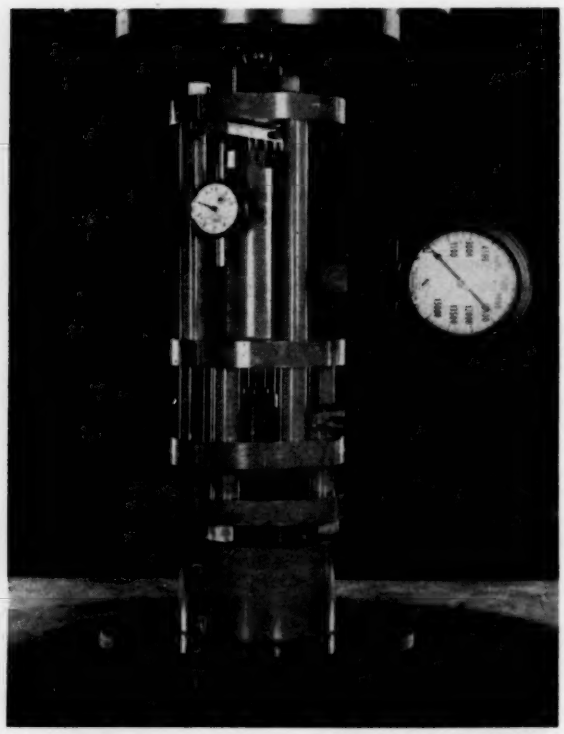


FIG. 1(*b*) APPARATUS FOR CONDUCTING STRENGTH TESTS ON ROCK UNDER HYDROSTATIC PRESSURE

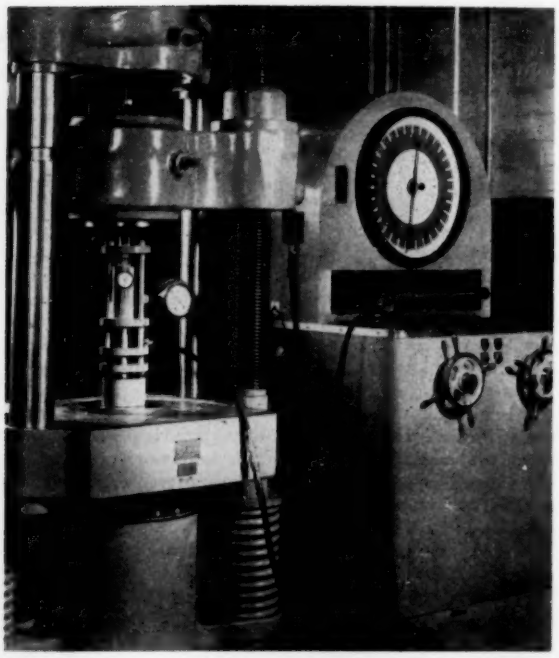


FIG. 1(*c*) STRENGTH-TEST APPARATUS IN PLACE IN BALDWIN TENSILE TESTING MACHINE
(Equipment is arranged for compressive tests.)

sign used by Bridgman (2), and later modified by Griggs (3–6) and Handin (7). Much lower confining pressures permitted the use of simpler seals than formerly had been applied.

The equipment was designed to accommodate a $^1/_2$-in-diam by 1-in-long specimen up to confining pressures of 21,000 psi. The specimen, S, is located in a pressure chamber V, and surrounded by a light hydraulic oil which acts as the pressure-transmitting agent. Soft-brass sleeves assist in aligning the upper piston P_1, the specimen, and the support R. The portion of the sleeves fitting over the specimen is slotted so that little restraint to lateral deformation of the specimen is allowed, other than the shear restraint on the face of the specimen. A piece of plastic tubing fitted over the specimen assembly protects it from contact with the confining liquid. In tests where the liquid is allowed access to the specimen, holes are cut into the tubing portion immediately surrounding the specimen.

Pistons P_1 and P_2, which are of equal diameters, are yoked together by three tie rods Y and plates L_1 and L_2. Piston P_2 is forced into the pressure cavity by a small hydraulic jack J, thus increasing the confining pressure around the specimen. The confining liquid has free access to both pistons, and the specimen is exposed to the same pressure in all directions. After attaining the desired confining pressure, load is applied to the specimen by compressive loading of plates L_1 and L_2. The confining pressure remains constant during the test, as pistons P_1 and P_2 remain a constant distance apart.

Sealing of the pressure cavity is accomplished by use of neoprene O-ring seals and closely ground fits between pistons, threaded plug T, and the pressure chamber.

Deformation measurements are taken by a dial indicator measuring total deformation of the specimen assembly, which includes piston P_1, the specimen, and support R. Subtraction of the extraneous deformations is accomplished by use of a steel specimen loaded within the elastic range at various confining pressures.

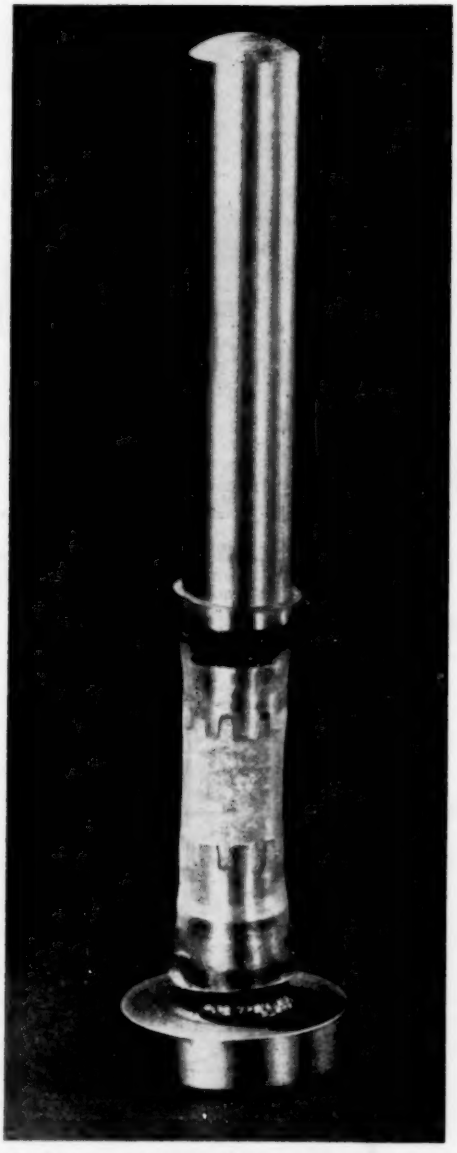


FIG. 1(*d*) SPECIMEN ASSEMBLY SHOWING UPPER PISTON, BRASS SLEEVES, SUPPORT, AND PLASTIC JACKET

Fig. 1(*c*) shows the pressure-test equipment in place in the Baldwin tensile test machine used for these tests.

SPECIMENS

Specimens were obtained from quarried rocks or cores which were taken from various well locations. The longitudinal axis of all specimens from cores was assumed to be parallel to the longitudinal axis of the core. Orientations of specimens from quarried samples were not known.

Thin sections from the various specimens were examined microscopically, together with corresponding hand samples. Resulting petrographic descriptions are presented in the Appendix.

The majority of specimens tested were in the air-dry condition; that is, no effort was made to control the moisture content of the specimens. The practice of jacketing the specimens followed that of Griggs who showed that the strength under high confining pressure is much greater when the specimen is jacketed than when the liquid is free to enter the pores. If the formation in question is relatively impermeable so that the rock-bit teeth advance ahead of the moisture penetration, then the condition of a dry, jacketed specimen appears to be the logical procedure. On the other hand, if a formation is permeable, then the condition of a saturated, unjacketed specimen is more logical. Several specimens were tested in the latter condition.

TEST RESULTS

Results of tests are presented as "stress-strain" curves. In those cases where all tests at a given confining pressure were consistent, an average curve is presented for the particular condition. When single tests deviated from the average determined by a group of other tests made under the same conditions, the stress-strain curve for the deviating test also was plotted.

The familiar stress-strain designation is employed in referring to the curves. The ordinate, however, is actually differential stress. The total axial stress is the differential stress plus the confining pressure. In determining the differential stress, the assumptions of constant volume and of uniform lateral deformation introduce an error when the specimen flows during the test, and acquires a barrel-like shape. However, this method of calculating stress is more accurate than the conventional method based upon the original cross section.

Appendix 2 gives a tabulation of the types of failures which occurred at various confining pressures. Measurements were made of the angles which the failure planes made with the longitudinal or compression axis of the specimen. In many instances, these measurements are inaccurate, particularly when the specimen failed by splitting longitudinally with only very small shear surfaces at the ends of the longitudinal splits. If the specimen remained intact, measurements were made of the angles outlined by a network of lines on the specimen surface. Griggs (3) states that these lines, which he takes to be surface expressions of internal shear surfaces, are analogous to Luders lines in metals.

The measured angles of fractured surfaces may or may not represent shear planes, as the failure planes, in many instances, appeared to be comprised of minute integrated areas of alternate "shear" and "tension" (splitting) type failures.

Failures at atmospheric pressure were generally of the tension type, occurring parallel to the direction of compression, with shear wedges at the ends of the specimen. Griggs (3) also had observed this type failure and developed the following hypothesis: Before rupture, a small amount of plastic deformation occurs, accompanied by development of Luders lines, which seem to be the external expression of internal shear surfaces. Where these surfaces intersect, a wedge is formed, with further movement along the shear surfaces causing development of tension at the point of the wedge. In response to the lateral wedge force, tension cracks develop. Griggs had made some tests in which he found that Luders lines actually developed and spread before the tension-type failures occurred.

In some cases, the fracture face was composed of failures at two angles. This usually occurred when the specimens were tested under confining pressures. The major portion of the failure (major failure plane) generally was oriented more nearly parallel to the compression direction than was the smaller plane, which was located near the end of the specimen. As confining pressures increased, the major failure angles generally increased. These observations are consistent with those of Griggs, who concluded that tension failures are inhibited by increased confining pressure. In fact, in examining his hypothesis of tension failures,

(*a*) Confining press. = 0 psi
Max. diff. stress = 6000 psi
Axial strain = 0.5 per cent

(*b*) Confining press. = 2500 psi
Max. diff. stress = 16,000 psi
Axial strain = 20.0 per cent

(*c*) Confining press. = 5000 psi
Max. diff. stress = 21,500 psi
Axial strain = 24.0 per cent

(*e*) Confining press. = 10,000 psi
Max. diff. stress = 30,000 psi
Axial strain = 21.0 per cent

(*d*) Confining press. = 7500 psi
Max. diff. stress = 27,000 psi
Axial strain = 24.5 per cent

FIG. 2 ANHYDRITE SPECIMENS AFTER COMPRESSION AT VARIOUS CONFINING PRESSURES—SHOWING TRANSITION FROM BRITTLE TO DUCTILE-TYPE BEHAVIOR

one notes that the effective tensile force at the point of the wedge is decreased by increased confining pressure. Thus, fracture planes change from the predominant tensile-type failure to failures consisting of more shear as confining pressures increase.

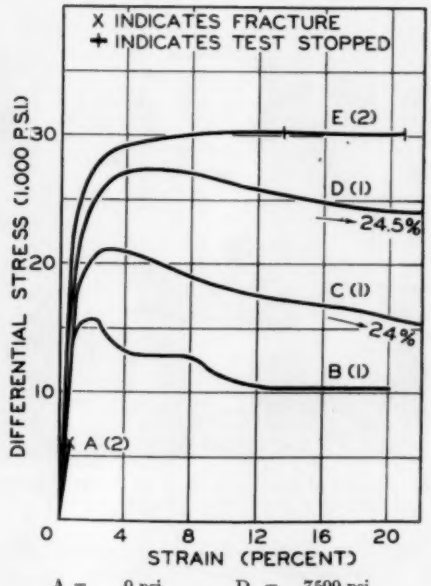
As noted earlier a network of lines was generally visible on the surface of specimens which deformed in a ductile manner. One of these (shale) was broken after removal from the test apparatus. The surface lines were measured at 38 deg to the compression axis, but the specimen broke at 32 deg. It is therefore not certain that the surface lines are indicative of planes upon which failure would occur if axial deformation were continued.

Generally, a marked increase in compressive strength with increasing confining pressure was noted in all formations. The only exception was in one group of sandy-shale specimens tested under conditions allowing exposure of the specimen to the confining liquid.

Anhydrite. At atmospheric pressure, anhydrite specimens were brittle. At pressures of 2500 psi and greater, the specimens behaved in a ductile fashion, Figs. 2 and 3. The apparent transition is in the pressure range 0 to 2500 psi. At 10,000 psi pressure, the strength of this material had increased to 30,000 psi from 6000 psi at atmospheric pressure—an increase of 400 per cent. In one specimen, Fig. 2(*e*), a strain of 21 per cent was achieved without loss of cohesion.

Austin Limestone. At atmospheric pressures, Austin limestone specimens failed at very low values of differential stress, Fig. 4. One specimen appeared to flow, but examination showed that it had split longitudinally.

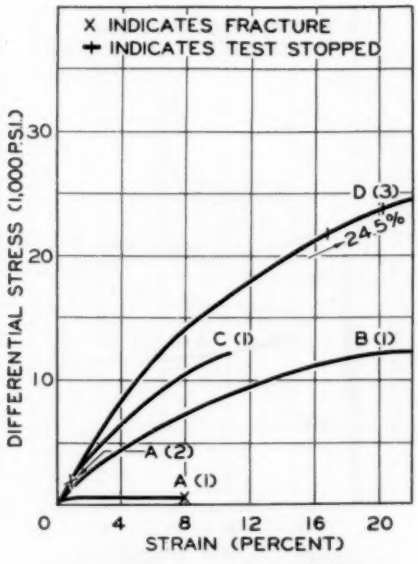
Failures at pressures of 5000 and 10,000 psi appeared to be by progressive crumbling, generally beginning near the center of the specimen. Apparently, the specimens acquired a barrel-like shape prior to crumbling.



A = 0 psi
B = 2500 psi
C = 5000 psi
D = 7500 psi
E = 10,000 psi

FIG. 3 STRESS-STRAIN CURVES OF ANHYDRITE TESTED IN COMPRESSION AT VARIOUS CONFINING PRESSURES

The curves show that this material has no distinct yield strength, but begins to flow immediately upon application of a differential stress.



A = 0 psi C = 7500 psi
B = 5000 psi D = 10,000 psi

FIG. 4 STRESS-STRAIN CURVES OF AUSTIN LIMESTONE TESTED IN COMPRESSION AT VARIOUS CONFINING PRESSURES

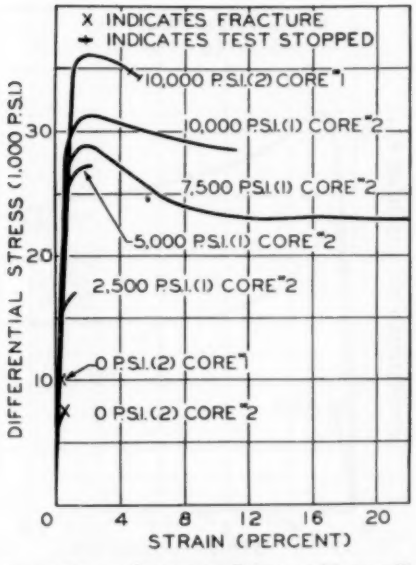


FIG. 5 STRESS-STRAIN CURVES OF CARTHAGE MARBLE TESTED IN COMPRESSION AT VARIOUS CONFINING PRESSURES

At the higher confining pressures, it was not possible to ascertain the points at which failures occurred.

Carthage Marble. Two different samples of this material were tested. Sufficient variation in the results was noted so that results for each were plotted in Fig. 5. Both samples were brittle at atmospheric pressure, with some longitudinal splitting evident. Both samples were ductile at 10,000-psi confining pressure, and one specimen from one sample was ductile at 7500-psi confining pressure.

A comparison of results on this formation may be made with those obtained by Handin and reported in an unpublished paper. Handin's tests were on a group of limestones and on Yule marble. All of these are relatively low strength formations at atmospheric pressure. At 1000 and 2000 atmospheres the formations which he tested were ductile and experienced large axial deformations without fracturing. In general, the results of his tests were consistent with those of the author.

Sandy Shale. This formation was tested under four different sets of conditions in an attempt to determine the effect of saturation by water upon jacketed and nonjacketed specimens. Formations may exist in the earth in various conditions. Some may be dry, others wet; and still others may be saturated with gas under pressure. Any one of the four conditions under which these tests were made may thus represent a condition existing in the actual formation.

Specimens which were saturated by water prior to testing gained about four per cent by weight during saturation at atmospheric pressure.

At atmospheric pressure specimens which were dry and jacketed were brittle, Fig. 7(*a*). At 5000 and 10,000-psi confining pressure, all specimens in this group were ductile. Fig. 6(*a*) shows a typical dry, jacketed specimen after having been deformed at 10,000-psi confining pressure.

Stress-deformation characteristics of specimens saturated prior to testing and jacketed during the test were similar to those of specimens which were dry and jacketed [see Fig. 7(*b*), Note 1].

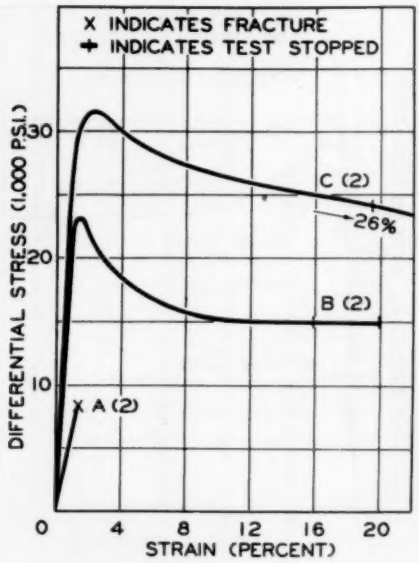
(*a*) Dry sandy shale
Confining press. = 10,000 psi
Max. diff. stress = 32,500 psi
Axial strain = 26.0 per cent

(*b*) Saturated sandy shale
Confining press. = 10,000 psi
Max. diff. stress = 22,000 psi
Axial strain = 18.5 per cent

FIG. 6 JACKETED SANDY SHALE SPECIMENS AFTER COMPRESSION AT 10,000-PSI CONFINING PRESSURE

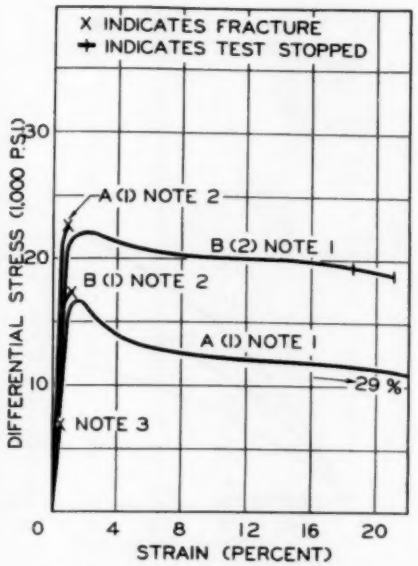
At 5000 and 10,000 psi ductile behavior again was noted. However, the differential stress which the formation could withstand before flow occurred seemed to be lowered appreciably as a result of saturation.

Behavior in the saturated, jacketed specimens was greatly different from those in the dry, jacketed specimens. The dry specimens barrelled at 10,000-psi confining pressure and remained intact even though deformed axially up to 26 per cent of the original length. Saturated specimens, on the other hand, broke into five or six parallel disks of approximately equal thicknesses on planes perpendicular to the compression direction. Figs. 6(*a*) and (*b*) show typical specimens in each group. A probable explanation of of this disk-type failure is that, upon compression, the pressure in the formation pores builds up to a large value because of the presence of the fluid. Upon release of the axial load and the con-



A = 0 psi B = 5000 psi
C = 10,000 psi

FIG. 7(a) STRESS-STRAIN CURVES OF SANDY SHALE TESTED IN COMPRESSION AT VARIOUS CONFINING PRESSURES. SPECIMENS WERE DRY AND JACKETED



A = 5000 psi B = 10,000 psi

FIG. 7(b) STRESS-STRAIN CURVES OF SANDY SHALE TESTED IN COMPRESSION AT VARIOUS CONFINING PRESSURES

(Note: 1—Saturated prior to test; jacketed. 2—Dry prior to test; not jacketed. 3—Saturated prior to test; not jacketed. Includes two specimens each at 0, 5000, and 10,000-psi confining pressure.)

(a) Confining press. = 0 psi
Max. diff. stress = 11,000 psi
Axial strain = 0.7 per cent

(b) Confining press. = 7500 psi
Max. diff. stress = 33,000 psi
Axial strain = 27.0 per cent

(c) Confining press. = 19,000 psi
Max. diff. stress = 38,000 psi
Axial strain = 19.5 per cent

FIG. 8 COARSE CHICO LIMESTONE SPECIMENS AFTER COMPRESSION AT VARIOUS CONFINING PRESSURES—SHOWING BRITTLE-TO-DUCTILE TRANSITION

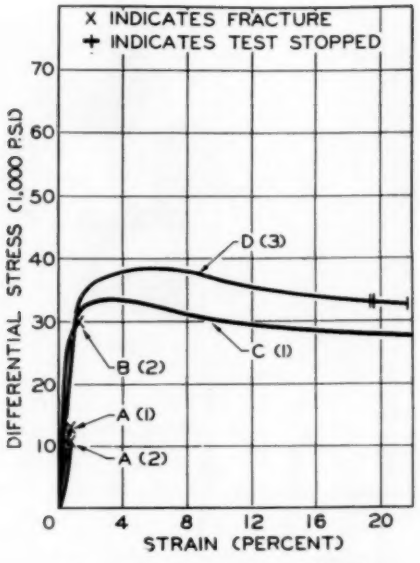
fining pressure, the internal fluid pressure cannot be immediately relieved because of the finite permeability of the formation. The resultant tensile force is sufficient to cause failure to occur parallel to the bedding planes, or perpendicular to the compression direction. Similar results were experienced by Griggs and Bell (8) in tests on quartz crystals.

All specimens, dry and saturated, which were exposed to the confining liquid were brittle. Those which were dry prior to testing [Fig. 7(*b*), Note 2] appeared to be inconsistent in behavior as the specimen tested at 5000-psi confining pressure supported a greater load than the one tested at 10,000-psi confining pressure.

Saturated specimens exposed to the confining liquid showed no increase in strength with increases in confining pressure [Fig. 7(*b*), Note 3]. In fact, the average strength, including all specimens at all confining pressures, was slightly below that of a dry, jacketed specimen at atmospheric pressure.

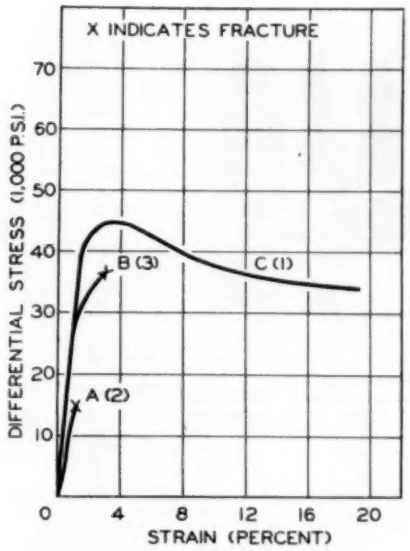
Coarse-Grained Chico Limestone. Results of tests on coarse-grained Chico limestone are shown in Figs. 8 and 9. This material is one which failed in typically brittle fashion at atmospheric pressure and 2500-psi confining pressure, but which was ductile at 7500 and 10,000-psi confining pressure. Axial deformation of 27 per cent, Fig. 8(*b*), without loss of cohesion was achieved in one specimen tested at 7500 psi.

Shale. Again, at atmospheric pressure, the rock was brittle. Fractures were predominantly of the longitudinal splitting, or tension type. Brittle behavior persisted at 10,000-psi confin-



A = 0 psi C = 7500 psi
B = 2500 psi D = 10,000 psi

FIG. 9 STRESS-STRAIN CURVES OF COARSE CHICO LIMESTONE TESTED IN COMPRESSION AT VARIOUS CONFINING PRESSURES



A = 0 psi B = 10,000 psi
C = 15,000 psi

FIG. 10 STRESS-STRAIN CURVES OF SHALE TESTED IN COMPRESSION AT VARIOUS CONFINING PRESSURES

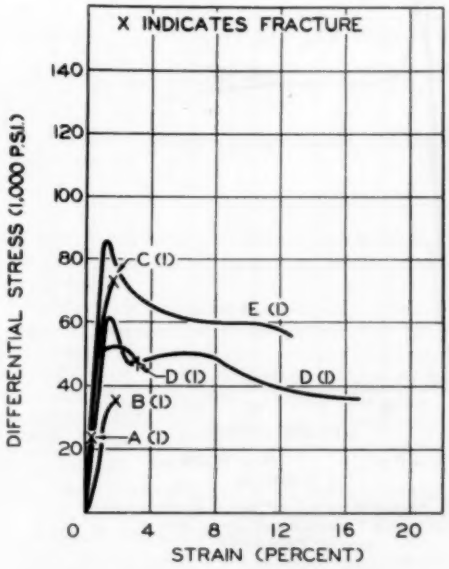
ing pressure but with a change in the character of fracture. Here, the specimens broke into only two pieces, with the failure planes at 30 and 38 deg.

Ductile behavior was seen at 15,000-psi confining pressure, with the specimen achieving 19.5 per cent axial strain without breaking. Surface lines were at 38 deg; however, after removal from the test apparatus, this specimen was easily broken at an angle of 32 deg.

D-1 Dolomite. Several inconsistencies appeared in the tests on D-1 dolomite, Fig. 11. The 5000-psi curve deviates from the elastic portion of the stress-strain curves established by all of the other tests. One specimen at 10,000 psi deformed axially 17.0 per cent and broke into many pieces. Another, at the same pressure, broke into many pieces after only 3 per cent axial compression. Still another inconsistency is in the results of the test at 7500-psi confining pressure, which show an ultimate strength exceeding that at 10,000-psi confining pressure.

At confining pressures of 10,000 and 15,000 psi, this formation appeared ductile in that it supported large loads after yielding. The appearances of the specimens after testing, however, suggest that neither 10,000 nor 15,000-psi confining pressure is sufficient to cause ductile behavior in the same sense as in previously discussed formations. Upon removal from the jacket, the specimens broke into many pieces. Another formation showing this same type of behavior was Wyoming red bed, failures of which are shown in Fig. 12.

Wyoming Red Bed. Inconsistent results at atmospheric pressure were obtained on this formation, Fig. 13. For two samples, the strength was 40,000 psi, while for another it was only 23,000 psi. The material was brittle at atmospheric pressure and at 10,000-psi confining pressure. At 15,000 psi, however, the specimen appears to have been at the borderline between ductile and brittle behavior. Although the stress-strain curve indicates flow in the material and the specimen assumed a barrel-like shape, fracture did occur. Fig. 12 shows the change in fracture appearance as confining pressure increased.



A = 0 psi D = 10,000 psi
B = 5000 psi E = 15,000 psi
C = 7500 psi

FIG. 11 STRESS-STRAIN CURVES OF D-1 DOLOMITE TESTED IN COMPRESSION AT VARIOUS CONFINING PRESSURES

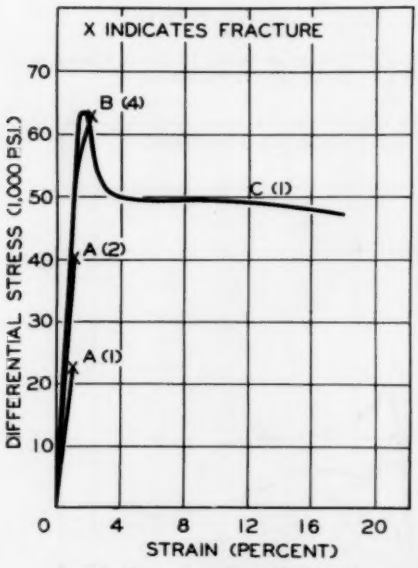
White Dolomite, Fine-Grained Chico Limestone, Rush Springs Sandstone, Knippa Basalt, Virginia Limestone. Results on these formations were all similar in that all failed in a brittle manner at atmospheric pressure, 10,000 psi, and 15,000 psi. Fracture characteristics were also similar—longitudinal splitting at atmospheric pressure and, in general, the breaking of the specimens into two

(*a*) Confining press. = 0 psi
Max. diff. stress = 23,000 psi
Axial strain = 0.9 per cent

(*b*) Confining press. = 10,000 psi
Max. diff. stress = 55,000 psi
Axial strain = 2.2 per cent

(*c*) Confining press. = 15,000 psi
Max. diff. stress = 65,000 psi
Axial strain = 17.0 per cent

FIG. 12 WYOMING RED BED SPECIMENS AFTER COMPRESSION AT VARIOUS CONFINING PRESSURES—SHOWING APPARENT BORDERLINE TRANSITION AT 15,000-PSI CONFINING PRESSURE



A = 0 psi B = 10,000 psi
C = 15,000 psi

FIG. 13 STRESS-STRAIN CURVES OF WYOMING RED BED TESTED IN COMPRESSION AT VARIOUS CONFINING PRESSURES

(*a*) Confining press. = 0 psi
Max. diff. stress = 10,500 psi
Axial strain = 0.8 per cent

(*b*) Confining press. = 10,000 psi
Max. diff. stress = 51,000 psi
Axial strain = 1.3 per cent

(*c*) Confining press. = 15,000 psi
Max. diff. stress = 52,000 psi
Axial strain = 1.2 per cent

FIG. 14 WHITE DOLOMITE SPECIMENS AFTER COMPRESSION AT VARIOUS CONFINING PRESSURES—SHOWING TYPICALLY BRITTLE BEHAVIOR AT ALL PRESSURES

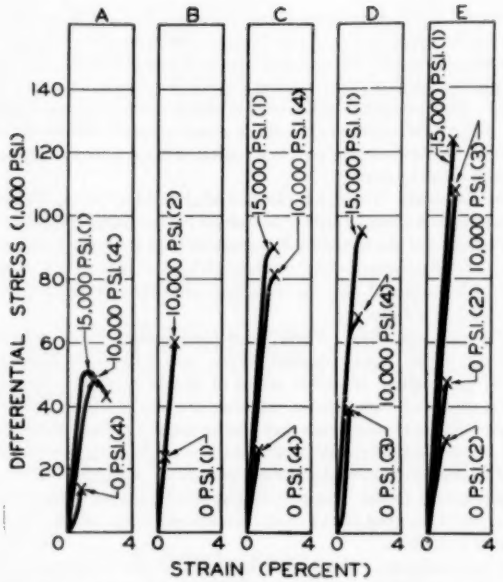
major pieces at 10,000 and 15,000 psi. Fig. 14 shows the typical failure behavior of these formations. The only formation which gave any evidence of ductile behavior was white dolomite at 15,000-psi confining pressure. Specimen appearance, Fig. 14(*c*), however, indicates the lack of large permanent deformation occurring prior to failure.

A wide range of rock strengths at atmospheric pressure is indicated in Fig. 15. It is noteworthy that these formations, which are greatly different in strength and composition, should behave in such similar fashion. At 15,000-psi confining pressure, the elastic axial strain before failure was almost invariably between 1.5 and 2.0 per cent.

It should be noted that the so-called limestone formations, Chico and Virginia, are actually dolomitic, as shown in Appendix I.

DISCUSSION

A summary of the effect of hydrostatic confining pressure upon the compressive strength of the various formations is shown in



A—White Dolomite
B—Fine Chico Limestone
C—Rush Springs Sandstone
D—Knippa Basalt
E—Virginia Limestone

FIG. 15 STRESS-STRAIN CURVES OF ROCK SAMPLES TESTED IN COMPRESSION AT VARIOUS CONFINING PRESSURES, COMPARING FORMATIONS HAVING BRITTLE CHARACTERISTICS TO 15,000-PSI CONFINING PRESSURE

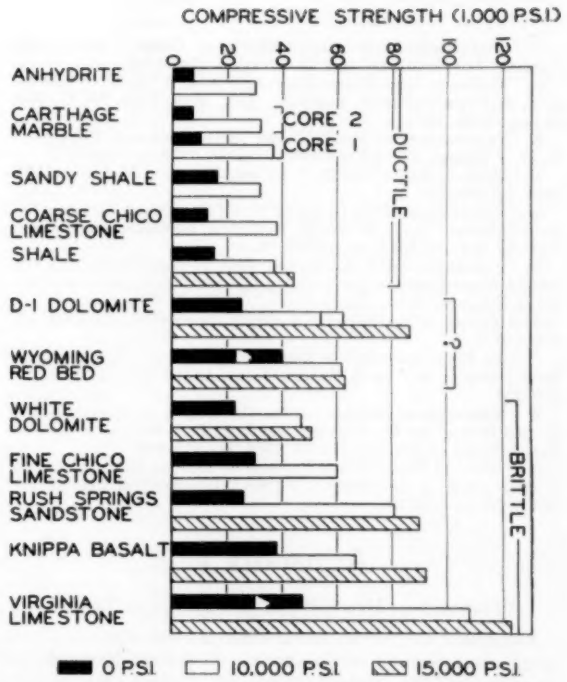


FIG. 16 BAR GRAPH SUMMARIZING RESULTS OF COMPRESSION TESTS AT 0, 10,000, AND 15,000-PSI CONFINING PRESSURES

Fig. 16. Formations have been grouped according to their behavior—ductile, borderline, or brittle—for the confining-pressure range 0 to 15,000 psi.

One fact is readily apparent. All formations which experienced a brittle-to-ductile transition have relatively low compressive strengths at atmospheric pressure. When these formations are exposed to hydrostatic pressures of 10,000 psi (or lower, in some cases) considerable deformation must take place before fracture occurs. To achieve fracture in such a formation in actual practice, it appears that preceding flow must be an important factor.

These low-strength formations are indeed best drilled when using rock bits which have long, widely spaced teeth and which provide a tearing-gouging action.

On the other hand, the higher-strength formations, which remain brittle up to 15,000 psi confining pressure, are best drilled using heavy-weight practice and rock bits providing a crushing-rolling action.

Another fact of importance is illustrated in Fig. 16. The compressive strength of all formations was increased considerably with increasing confining pressure, including the stronger formations such as Knippa basalt and Virginia limestone. This appears to support the contention that hydrostatic pressures in the range investigated have an important effect upon drilling rates, even in those formations where changes in failure characteristics were not obtained. Perhaps even more valid support of this contention is found in results of tests described by Murray and Cunningham (9), in which cores 3-in. diam were exposed to various confining pressures and drilled with a $1^1/_4$-in-diam bit. When drilling on Ellenberger dolomite, which is similar to Virginia limestone in strength and composition, a decrease of 50 per cent in drilling rate resulted as confining pressure was increased from atmospheric to 5000 psi, when using a 2000-lb bit load. Murray and Cunningham also cite results of field tests in which hydrostatic pressure affected drilling rates approximately the same as confining pressure in the laboratory.

CONCLUSIONS

1 Increases in compressive strength with increasing confining pressure were obtained on all formations tested in the dry, jacketed condition.

2 Brittle behavior was obtained at atmospheric pressure on all formations tested in the dry, jacketed condition.

3 A brittle-to-ductile transition was obtained in the confining-pressure range 0 to 15,000 psi on anhydrite, Carthage marble, sandy shale, coarse Chico limestone, and shale specimens.

4 Formations which appeared to be near the brittle-to-ductile transition at 15,000-psi confining pressure were D-1 dolomite and Wyoming red bed.

5 Brittle behavior was obtained at all confining pressures to 15,000 psi on White dolomite, fine Chico limestone, Rush Springs sandstone, Knippa basalt, and Virginia limestone.

6 Fracturing of rock exposed to hydrostatic pressure in many cases is dependent upon the deformation as well as the stress which is imposed upon the formation. Thus, in formations which become ductile under pressure, large amounts of flow must occur before fracture is achieved, and ultimate removal of the rock is accomplished.

ACKNOWLEDGMENT

The author wishes to thank the Hughes Tool Company for permission to publish this paper. Also, the author is grateful to Dr. John Rogers, of the Rice Institute, for the petrographic descriptions, and to Dr. John Handin, Shell Development Company, Houston, Texas, and Mr. H. B. Woods, Hughes Tool Company, for their constructive criticisms of the paper.

BIBLIOGRAPHY

1 "Festigkeitsversuche unter allseitigem Druck," by T. von Karman, *Zeitschrift VDI*, vol. 55, 1911, pp. 1749-1757.

2 "Studies in Large Plastic Flow and Fracture," by P. W. Bridgman, McGraw-Hill Book Company, Inc., New York, N. Y., first edition, 1952, 362 pp.

3 "Deformation of Rocks Under High Confining Pressures," by D. T. Griggs, *Journal of Geology*, vol. 44, 1936, pp. 541–577.

4 "Creep of Rocks," by D. T. Griggs, *Journal of Geology*, vol. 47, 1939, pp. 225–251.

5 "Experimental Flow of Rocks Under Conditions Favoring Recrystallization," by D. T. Griggs, *Bulletin of the Geological Society of America*, vol. 51, 1940, pp. 1001–1022.

6 "Deformation of Yule Marble: Part I—Compression and Extension Experiments on Dry Yule Marble at 10,000 Atmospheres Confining Pressure, Room Temperature," by D. T. Griggs and W. B. Miller, *Bulletin of the Geological Society of America*, vol. 62, 1951, pp. 853–862.

7 "An Application of High Pressure in Geophysics: Experimental Rock Deformation," by J. Handin, Trans. ASME, vol. 75, 1953, pp. 315–324.

8 "Experiments Bearing on the Orientation of Quartz in Deformed Rocks," by D. T. Griggs and J. F. Bell, *Bulletin of the Geological Society of America*, vol. 62, 1938, pp. 853–862.

9 "Effect of Mud Column Pressure on Drilling Rates," by A. S. Murray and R. A. Cunningham, AIME Petroleum Transactions, vol. 204, 1955, pp. 196–204.

Appendix 1

Petrographic Descriptions of Formations Tested

Anhydrite. Cores from Leduc Field, Alberta, Canada. Rectangular anhydrite grains ranging from 0.1 to 2 mm long make up approximately two thirds of the sample. Cutting the anhydrite and filling the interstices between the anhydrite grains is fine-grained, fibrous, gypsum which comprises another one third of the rock but varies markedly in quantity from place to place. One per cent of the rock consists of irregular grains of carbonate. The sample is massive.

Austin Limestone. Rock quarried near Austin, Tex. Approximately 60 per cent of the sample is composed of recrystallized foraminifera with an average length of 0.2 mm and a width of 0.1 mm. The foraminifera are completely recrystallized to a mass of fine-grained calcite and are commonly slightly stained by ferruginous material. The matrix, which comprises 40 per cent of the rock, consists of interlocking calcite grains with an average diameter of 0.1 mm. Less than 0.5 per cent of fine-grained silica is scattered through the rock. Except for a moderately developed parallelism (perpendicular to the length of the core) of some of the foraminifera, the rock is massive.

Carthage Marble. Rock quarried at Carthage, Mo. Approximately 40 per cent of the rock consists of foraminiferal and other recrystallized, fine or coarse-grained carbonate fossil fragments. These fragments are scattered randomly throughout a matrix of irregular, interlocking, 1-mm carbonate (calcite ?) grains which comprises 60 per cent of the sample. The sample is massive.

Sandy Shale. From cores from Leduc Field, Alberta, Canada. The rock is a moderately sorted sandstone in which the sand grains are subangular, irregular, and have an average diameter of 0.2 mm. Sand grains make up 70 per cent of the rock and consist of: Quartz and chert, 70 per cent; feldspar (mainly plagioclase), 20 per cent; and volcanic fragments, 10 per cent. The matrix, which comprises 30 per cent of the rock, is about one half carbonate and one half silica. The sample shows a very rough bedding perpendicular to the core.

Coarse-Grained Chico Limestone. Rock quarried at Chico, Tex. About one third of the rock consists of an unsorted mass of fine-to-coarse foraminifera, and other shell fragments. These clasts are imbedded in a matrix of very fine-grained limestone which comprises the remainder of the rock. The rock is massive.

Shale. From cores from Duval County, Tex. The sample is essentially a silty shale. Approximately 40 per cent of the sample consists of angular, 0.02-mm grains of quartz and feldspar oriented roughly parallel to the bedding (which is perpendicular to the core). The remainder of the rock is mainly a carbonaceous, illitic (?), slightly ferruginous clay which shows a rough unit extinction parallel to the bedding. A few scattered tiny grains of opaque material may be pyrite.

D-1 Dolomite. From cores from Leduc Field, Alberta, Canada. The sample is almost entirely carbonate (either calcite or dolomite or a mixture of the two) but does contain rare, irregular patches of fine-grained gypsum. Approximately 95 per cent of the carbonate is very fine grained and the remainder consists of irregular grains up to 0.2 mm diam.

Wyoming Red Bed. From cores from Sussex Field, Johnson County, Wyo. Approximately 30 per cent of the rock consists of 0.05-mm, angular, irregular grains of quartz and some feldspar; a few grains of dark opaque minerals are also present. Another 65 per cent is a ferruginous clay containing some streaks of darker red material which roughly parallel the crude bedding (perpendicular to the core) shown by the silt grains. A mineral of moderate birefringence forms irregular, fine-grained patches which make up about 5 per cent of the rock; the mineral is tentatively identified as anhydrite.

White Dolomite. From quarried rock, source unknown. The sample is most accurately classed as a serpentinized forsterite-diopside-calcite marble. About 70 per cent of the rock consists of interlocking 0.5-mm grains of calcite. Twenty per cent of the rock is composed of fine-grained fibrous serpentine which appears to replace the calcite and also cuts the 0.3-mm forsterite grains into islands of small grains surrounded by serpentine. Forsterite comprises about 5 per cent of the rock, and 0.3-mm diopside grains (also slightly altered to serpentine?) comprise another 5 per cent. The materials appear unoriented.

Fine-Grained Chico Limestone. Rock quarried at Chico, Tex. Euhedral carbonate (dolomite ?) rhombs with an average diameter of 0.01 mm form approximately 50 per cent of the sample. These rhombs are imbedded in a network of interlocking calcite grains which comprise 30 per cent of the rock and range up to 0.1-mm diam. Approximately 5 per cent of the rock consists of rounded, subspherical, 0.1 to 0.5-mm clasts of fine-grained, iron-stained limestone (probably altered foraminifera). The remaining 20 per cent of the rock is composed of well-rounded subspherical, 0.2-mm grains of quartz which show a rough bedding parallel to the long axis of the core.

Rush Springs Sandstone. Rock quarried at Cement, Okla. Well-sorted, subrounded, subspherical, 0.1 to 0.2-mm sand grains comprise about 60 per cent of this sample. Three fourths of these grains are quartz (and rare chert) and the remaining one fourth is potash feldspar and plagioclase. The cement, which comprises 40 per cent of the rock, consists of 0.1 to 0.2-mm grains of chemical carbonate (either calcite or dolomite). Rare tiny grains of pyrite are scattered through the sample. The rock is massive.

Knippa Basalt. Rock quarried at Knippa, Tex. The sample is a porphyritic basalt in which phenocrysts of enstatite and pigeonite (0.5 mm long) comprise about 30 per cent of the rock. The ground mass, which comprises 70 per cent of the rock, consists of clinopyroxene (0.02-mm grains, 20 per cent of groundmass), a very finely crystalline or glossy matrix (45 per cent of groundmass) and magnetite or other dark opaque mineral (15 per cent of groundmass). Both phenocrysts and matrix show a poorly developed flow banding at roughly 30 deg to the axis of the core.

Virginia Limestone. Rock quarried at Ripplemead, Va. Almost 100 per cent of the sample consists of equigranular, 0.05-mm

rhombs of carbonate; very slow effervescence in 6N HCl indicates that the carbonate is probably dolomite. Scattered through the rock are rare grains of course carbonate and fine-grained quartz and feldspar. One section parallel to the length of the core contains a veinlet of medium-grained, rutilated (?) quartz grains.

Appendix 2

Description of Failures

General Note: Angles given are measured with respect to the longitudinal axis of the specimen. In those cases where specimens remained intact after testing, angles are those made by surface lines measured with respect to the longitudinal axis. Angles are only approximate, as in most cases fractured surfaces were very irregular.

Formation	Confining pressure, psi	Failures	Angles, deg: Surface	Major	Wedge
Anhydrite					
	0	Some longitudinal splits	..	16–18	...
	0	Two major pieces	..	18	34
	2500	Barrelled; wedges at ends	32	.....	32
	5000	Barrelled	32	.....	.....
	7500	Barrelled	36	.....	.....
	10000	Barrelled	36	.....	.....
	10000	Barrelled	36	.....	.....
Austin limestone					
	0	Two major pieces	..	15	29
	0	Many pieces	..	0	.....
	0	Two major pieces	..	0	36
	5000	Crumbling	..	.....	.....
	7500	Slight end chipping; no major failure	..	.....	.....
	10000	Crumbling near both ends	..	.....	.....
		Center barrelled			
	10000	Crumbling near center	..	.....	.....
	10000	Crumbling near center	..	.....	.....
Carthage marble					
Core 1	0	Two major pieces; some longitudinal splits	..	26	.....
	0	Same	..	26	.....
	10000	Barrelled	26	.....	.....
	10000	Barrelled	36	.....	.....
Carthage marble (continued)					
Core 2	0	Two major pieces; some longitudinal splits	..	18	36
	0	Many pieces; mostly longitudinal splits	..	0	26
	2500	No fracture	..	.....	.....
	5000	No fracture	36	.....	.....
	7500	Barrelled	36	.....	.....
	10000	Barrelled	36	.....	.....
Sandy shale					
Jacketed, Dry	0	Two major pieces	..	22	.....
	0	Two major pieces	..	24	.....
	5000	Barrelled	37	.....	.....
	5000	Barrelled	35	.....	.....
	10000	Barrelled	37	.....	.....
	10000	Barrelled	45	.....	.....
Jacketed, Saturated	5000	Barrelled on one end; other end broke off in disk	41	90	.....
	10000	Six parallel disks	..	90	.....
	10000	Five parallel disks	..	90	.....
Non-jacketed, Dry	5000	Two major pieces; some longitudinal splits	..	38	.....
	10000	Two major pieces	..	48	.....
Non-jacketed, Saturated	0	Two major pieces; some crumbling	..	21	.....
	0	Two major pieces	..	21	.....
	5000	Two major pieces	..	21	.....
	5000	Four major pieces	..	26	.....
	10000	Two major pieces	..	25	.....
	10000	Two major pieces	..	22	38
Coarse-grained Chico limestone					
	0	Mostly longitudinal splits	..	.....	34
	0	Same	..	.....	28
	0	Same	..	.....	28–31
	2500	Two major pieces	..	29	.....
	2500	Two major pieces	..	26–32	.....
	7500	Barrelled	33	.....	.....
	10000	Barrelled	34	.....	.....

Formation	Confining pressure, psi	Failures	Angles, deg: Surface	Major	Wedge
Shale					
	0	Many pieces; mostly longitudinal splits	..	.....	17
	0	Same	..	.....	23–15
	10000	Two pieces	..	30	.....
	10000	No failure	..	.....	.....
	10000	Two pieces	..	38	.....
	15000	Barrelled	38	32[a]	.....
D-1 dolomite					
	0	Mostly longitudinal splits; many pieces	..	.....	25
	2500	Many pieces	..	.....	24
	5000	Two major pieces	..	24	34
	10000	Barrelled before failing in many pieces	..	29	.....
	10000	Many pieces	..	30	.....
	15000	Many pieces; barrelled before failure	..	29	.....
Wyoming red bed					
	0	Many pieces; mostly longitudinal splits	..	.....	24
	0	Same	..	.....	25–16
	0	Same	..	.....	23–26
	10000	Three major pieces	..	23–33	.....
	10000	Three major pieces; some longitudinal splits	..	25	.....
	10000	Two pieces	..	29	.....
	10000	Two pieces	..	21–33	.....
	15000	Barrelled on one end before fracture	35	.....	.....
White dolomite					
	0	Mostly longitudinal splitting	..	.....	29
	0	Same	..	.....	29
	0	Same	..	.....	32
	0	Same	..	.....	23
	10000	Two major pieces	..	32	38
	10000	Two major pieces	..	32	.....
	10000	Two major pieces	..	18	28
	10000	Two major pieces	..	27	30
	15000	Two major pieces	..	34	.....
Fine-grained Chico limestone					
	0	Many pieces; longitudinal splits	..	.....	24
	10000	Two major pieces	..	24	.....
	10000	Two major pieces	..	29	.....
Rush Springs sandstone					
	0	Many pieces; mostly longitudinal splits	..	.....	18
	0	Same	..	.....	18
	0	Same	..	.....	21
	0	Same	..	.....	18
	10000	Two pieces	..	21	41
	10000	Many pieces	..	35	.....
	10000	Three major pieces	..	30	.....
	10000	Two major pieces	..	28	.....
	15000	Two major pieces	..	28	.....
Knippa basalt					
	0	Many pieces; mostly longitudinal splits	..	.....	22
	0	Same	..	.....	22
	0	Same	..	.....	21
	10000	Two pieces	..	28	45
	10000	Two pieces	..	31	36
	10000	Two pieces	..	30	.....
	10000	Two major pieces	..	26	.....
	15000	Two pieces	..	21	.....
Virginia limestone					
	0	Many pieces; mostly longidinal splits	..	.....	35
	0	Same	..	.....	.....
	0	Same	..	.....	23
	0	Same	..	.....	23
	10000	Two major pieces	..	18	28
	10000	Two major pieces	..	18	41
	10000	Two major pieces	..	16	35
	15000	Two major pieces	..	21	24

[a] Broken after termination of test.

Discussion

JOHN HANDIN.[3] Since the pioneering work of the great engineer von Karman[4] nearly a half century ago, the field of experi-

[3] Shell Development Company, Houston, Texas.
[4] See author's Bibliography (1).

mental rock deformation under high confining pressure has been neglected by all but a very few investigators, most of whom are geologists or geophysicists. Notable exceptions are the studies of Bridgman[5] and of the Denver laboratories of the Bureau of Reclamation.[6] However, Bridgman's experiments were conducted mostly under extreme hydrostatic pressures (about 25,000 atm) far above those pertinent to petroleum engineering. The confining pressures of the Bureau's triaxial tests (a few hundred atmospheres) were usually too low to induce plasticity in ordinarily brittle rocks. Until recently the experiments inspired by a principally geological interest had been made under pressures of the order of 5000 to 10,000 atm. Therefore, the author has made a most important and welcome contribution to the fund of knowledge of the behavior of rocks under moderate confining pressures.

His results along with the recent work of Robertson[7] and of the writer[8] should be useful to the engineer concerned with the failure of rocks at moderate depths in the earth's crust. Indeed, the awakened interest of engineers in the mechanical properties of buried rocks is most gratifying. In the past, much time and effort have been devoted to improving the tools of well drilling, but the material being penetrated, namely, the rocks, has been largely ignored. The oil-tool industry is to be congratulated for its support of fundamental research into the drillability of rocks under confining pressure.

The results of the author, Robertson, and the writer from triaxial compression tests on jacketed rock specimens are all consistent. The effects of increased confining pressure are increases of both ultimate strength and ductility in some measure depending upon the structure and composition of a particular rock. Both effects are adverse to drilling by any method which requires that the rock be broken. The energy needed to fracture a rock is proportional to the area under the stress-strain curve carried to rupture. Under confining pressures appropriate to oil-well drilling depths, this area may increase by two or three orders of magnitude over its value at atmospheric pressure.

Probably everyone will agree that the data of triaxial tests cannot be applied directly to the problem of rotary drilling at the present time. As the author has recognized, the state of stress in the rock at the bottom of a bore hole under the action of a bit must be exceedingly complex. In the first place, the stress even in the undrilled rock would seldom be hydrostatic since wells are commonly drilled in tectonically unstable regions, the usual habitat of oil. Furthermore, the pre-existing stress pattern would be perturbed by the bore hole itself. Finally, the bit does not apply a simple compressive load to the rock.

Even if these complications are discounted, and a uniform state of stress is assumed in the rock just below the bit, there still remains the problem of fluid pressures. In the rocks of the shallow crust, which are seldom dry, there is normally a fluid pore pressure about equal to the head of a column of water to the surface; that is, about 40 per cent of the confining pressure exerted by the weight of superincumbent rock. If the drilling fluid does not invade the formation, there is, in addition, a fluid pressure on the bottom of the hole which in general differs from the partial pore pressure. The jacketed specimen represents the unlikely situation in which a formation is initially quite dry and is then uninvaded by the drilling fluid. The unjacketed specimen simulates a porous formation which has been invaded so that the confining pressure equals the weight per unit area of the column of drilling fluid.

The author shows that both strength and ductility of an unjacketed rock are virtually uninfluenced by hydrostatic pressure, in sharp contrast with the changes effected in jacketed samples The actual conditions in the earth are probably intermediate between the two extremes, and what is needed is a series of experiments on jacketed specimens in which there are independently controlled partial-pore pressures. The preliminary results of such experiments by H. Heard (personal communication, June, 1956) on Solenhofen limestone suggest that the influence of pore pressure on strength and ductility is small until that pressure closely approaches the confining pressure.

In addition to the incompletely investigated factor of fluid content, at least two other variables, temperature and strain rate, must be taken into account.

In the crust, some temperature gradient is of course always associated with the pressure gradient, and while temperatures of the order of 100–200 C do not affect the pertinent rock properties as greatly as the corresponding pressures, enough work has been done to show that the influence of temperature cannot be neglected. There is usually, but not always, a decrease in strength and an increase in ductility. A striking example is limestone whose yield stress at 20 C exceeds that at 150 C by about 30 per cent under 5000-atm confining pressure.[9]

The effect of strain rate is largely unexplored, at least in combination with confining pressure. Robertson[7] made one triaxial experiment on jacketed limestone subjected to nearly impact loading. The yield stress was discovered to be about 20 per cent greater than for loading at the ordinary strain rate of the order of 10^{-3} per sec. By analogy with metals one might expect both strength and ductility of rocks to increase with strain rate at least for speeds associated with rotary drilling. However, in the present state of knowledge this is pure speculation.

These remarks are not intended to imply that the author's work and that of others on the influence of confining pressure alone is without significance to the drilling problem. Indeed, of one thing we may be sure even in our ignorance of the details, the properties of rocks as measured by ordinary crushing, tensile, or bit performance tests at atmospheric pressure are not the properties of the rock at the bottom of a 10,000-ft well. This discussion is intended not to discourage the use by engineers of the data now available, incomplete though these data may be, but rather to encourage further research in rock deformation. Experiments are needed in which the natural environment of a buried rock is simulated as realistically as possible. The results of well-conceived experiments can hardly fail to aid in the improvement of existing drilling methods and perhaps in the development of new techniques.

J. J. W. Rogers.[10] In commenting upon this very excellent paper, the writer feels that the author has provided structural geologists with much needed information on the laboratory properties of a number of common rocks. Wherever possible, geologic information, drawn from field studies of naturally deformed rocks, should be correlated with laboratory data of this type. In this brief discussion, the writer would like to review the relation between the author's results and three common geological observations on deformed rocks.

[5] "Large Plastic Flow and Fracture," by P. W. Bridgman, McGraw-Hill Book Company, Inc., New York, N. Y., 1952.

[6] "Physical Properties of Some Typical Foundation Rocks," by G. G. Balmer, U. S. Bureau of Reclamation, Concrete Laboratory Report SP-39, 1953.

[7] "Experimental Study of the Strength of Rocks," by E. C. Robertson, *Bulletin of the Geological Society of America*, vol. 66, 1955, pp. 1275–1314.

[8] "Experimental Deformation of Sedimentary Rocks Under Confining Pressure. Part I—Tests at Room Temperature on Dry Samples," by John Handin and R. V. Hager, Jr., 1957 (in press).

[9] "Deformation of Yule Marble. Part II—Effects at 150° C," by David Griggs, F. J. Turner, Iris Borg, and John Sosoka, *Bulletin of the Geological Society of America*, vol. 62, 1951, pp. 1385–1406.

[10] Department of Geology, The Rice Institute, Houston, Texas.

One observation so universally made by geologists in the field that it has become a fundamental principle of structural geology is the distinction between two basic types of rocks—the competent and the incompetent. The relatively strong, competent rocks are capable of resisting differential stress until fracturing occurs. As observed in nature, these rocks, such as limestone, dolomite, sandstone, and igneous and metamorphic rocks, tend to fracture without previous flowage. In distinction, naturally occurring incompetent beds tend to flow without fracturing. Typical examples of incompetent rocks are shale and easily recrystallized materials such as gypsum and other saline deposits. The border between the competent and incompetent rocks is, of course, a gradational one, and the distinction loses significance at depth, where high temperature and pressure and the effects of metamorphism tend to make all rocks structurally similar. Nevertheless, in the outer portions of the crust, where pressures are in the range investigated by the author, the structural difference between competent and incompetent materials is clearly evident.

The question now arises as to whether his distinction between brittle and ductile rocks matches the geologic distinction of the competent and incompetent. In the present study, the brittle rocks were the white dolomite, fine Chico limestone, Rush Springs sandstone, Knippa basalt, and Virginia limestone. Petrographic examination confirms that three of these five rocks are carbonates, one is sandstone, and one is basalt. On the basis of field work, structural geologists would expect all of these to be competent and, therefore, brittle. Consequently, the correspondence between laboratory and field evidence appears to be very good.

Unfortunately, however, in the case of those rocks found to be ductile, the relation between field and laboratory evidence is not clear. The ductile rocks were anhydrite, Carthage marble, sandy shale, coarse Chico limestone, and shale. Of these rocks, the shale and anhydrite (gypsiferous) are expected to be incompetent and ductile. The sandy shale (a fine-grained sandstone), Carthage marble, and coarse Chico limestone, however, would normally be considered competent and brittle. For some reason, these three rocks have behaved in the laboratory in a manner unlike their expected behavior in the field. Naturally, it would be interesting to know the reasons for this discrepancy between field and laboratory data, for a knowledge of the conditions under which these rocks would deform, as brittle rather than ductile material would give geologists a better understanding of the conditions existing during deformation of the earth's surface. Presumably, conditions during geological deformation do not conform precisely to the conditions under which the author's experiments were conducted, but the exact difference between nature and laboratory is not apparent.

In addition to the basic principle of the competency of geologic formations, two other important concepts of structural geology may be discussed in connection with his work. One principle is that the presence of water generally weakens rocks and causes them to flow more easily. Another concept is that increasing hydrostatic pressure causes rocks to be less brittle and more ductile.

The effect of water in lowering the strength of rocks is generally confirmed by the author's experimental comparison of dry and saturated samples of the sandy shale. As expected by geologists, saturated samples deformed at a lower differential stress than dry samples.

Finally, his data also confirm geologic deductions to the effect that increasing hydrostatic pressure makes rocks tend to flow rather than to fracture. It has long been known that, in the deeper parts of the earth's crust, folding and flowage of rocks predominate over fracturing, even in the case of those rocks which are normally brittle near the surface. This tendency to flow, however, could be attributed equally well to an increase of temperature or an increase of pressure with depth. Obviously, higher temperatures make rocks more plastic. There has been some evidence, however, both geologic and experimental, that the pressure of overlying rocks would cause ductility at depth even in the absence of high temperature. The writer is happy to note that the author's work supports this conclusion.

In summary, it seems that his data are in keeping with a number of the conclusions reached by structural geologists working in the field. In a few cases (such as the ductile behavior of a marble and a limestone) the two lines of investigation have reached different conclusions, presumably because the conditions under which rocks were deformed naturally were different from the experimental conditions. It is to be hoped that additional work eventually will determine the nature of these differences.

Author's Closure

The author wishes to express his appreciation to Dr. Handin and Dr. Rogers for their interest in and their comments on this paper.

As Dr. Handin has observed, the data of triaxial tests cannot be applied *directly* to the problem of rotary drilling at the present time and experiments are needed in which the natural environment of a buried rock are simulated as realistically as possible.

The results of some tests in which simulated field drilling conditions were employed have been reported by Murray and Cunningham.[11] A $1^1/_4$-in. diam two-cone microbit was used to drill upon 3-in. diam cores. The face being drilled was exposed to hydrostatic pressure simulating fluid or "mud" column pressure. These tests, while neglecting the effects of fluid pore pressure and, to a large extent, overburden pressure, showed that fluid column pressure has a definite influence upon drilling rate.

Tests in which the effects of pore pressure and overburden pressure, as well as the type of drilling fluid being used (high-water-loss versus low-water-loss muds), will be conducted in the near future by Cunningham.[12] As before, a $1^1/_4$-in. diam bit will be used to drill upon 3-in. diam cores. Features of the new equipment are:

1 The complicated stress condition at the bottom of the hole will be simulated by the bottom hole pattern formed by the microbit.

2 A fluid pressure up to 10,000 psi on the top of the core will simulate the fluid column pressure.

3 A fluid pressure up to 10,000 psi will be introduced into the formation pores from the bottom of the core.

4 A lateral fluid pressure up to 15,000 psi on the outer cylindrical surface of the core will simulate overburden pressure. This fluid will be sealed from the core by a plastic or soft metal jacket.

In commenting upon Dr. Roger's observation of the discrepancies between field and laboratory data in the cases of the sandy shale, Carthage marble, and coarse Chico limestone, the author can state only that it is extremely unlikely that conditions during geological deformation conform with laboratory conditions in the author's tests. Impending tests (discussed by Dr. Handin) in which variables such as pore pressures, temperatures, and strain rates are considered, may offer evidence sufficient to resolve these discrepancies.

---

[11] "Effect of Mud Column Pressure on Drilling Rates," by A. S. Murray and R. A. Cunningham, AIME Petroleum Transactions, vol. 204, 1955, pp. 196–204.

[12] Per verbal communication.

Prevention of Destructive Engine Failure by Spectrographic Analysis of Crankcase Oils

By R. E. LINNARD,[1] C. B. THRELKELD,[1] and R. T. BLADES,[2] BARTLESVILLE, OKLA.

Crankcase oils of stationary engines operating in relatively constant service normally contain equilibrium but minute amounts of certain metals indicative of normal engine wear. The occurrence of abnormal wear or corrosion conditions in an engine is reflected in concentration of one or more of the key wear metals in the crankcase oil. These trends in metals content offer important and early clues concerning the abnormal events which are occurring and suggest the need for corrective action usually far in advance of destructive engine failure. Utilization of data from emission spectrographic analysis of crankcase oils has been found to have practical use in the surveillance of pipeline-engine operations, particularly where the performance of copper-lead bearings has been in question. The application of this type of used-oil analysis as a practical aid in the prevention of destructive engine failures is reviewed, the interpretation of data is discussed, and examples of the relationship between increasing key wear-metals content and observed engine conditions are cited.

Introduction and Background

THE old adage "an ounce of prevention is worth a pound of cure" was never more true than when considered from the viewpoint of preventing destructive failures in today's modern and efficient, but expensive, pipeline pumping engines. When failures are encountered, to the direct costs of replacement parts and labor must be added the indirect expense of down time, lack of engine availability, and inefficient operations. Therefore, it behooves every power-plant operator to consider all sound techniques available which afford some indication of the transient conditions occurring in malfunctioning engines as a means of detecting areas of distress of certain components whose complete failure might lead to more serious engine damage.

Toward this end, it is the purpose of this paper to bring to the attention of the industry a used-oil-analysis technique found to be a reliable indicator of abnormal engine conditions where the lack of detection would lead to destructive failures in many instances. In the following, the application of the technique of emission spectrographic examination of used-oil samples taken from stationary engines is described as a practical aid in controlling engine operations of the pipeline system.

Essentially, the technique consists of determining the trace concentrations of certain key wear metals in samples of used oil taken from the engine crankcases at regular intervals. In engines operating normally these wear metals are present in minute but equilibrium amounts. Analyses which show increasing trends in concentration of one or more of these wear metals in the used oil offer early clues as to what sort of unusual conditions might be developing and suggest the need for engine inspection at a time usually far in advance of complete failure of the engine component in distress.

The utilization of emission spectrographic analyses of used crankcase oils as a means of finding out what events are taking place inside an engine is not new. Much of the credit for developing such techniques belongs to the railroad industry wherein the general scheme was applied on a practical basis by a number of the Class I railroads during their dieselization programs in the postwar years. Many papers have been published (1, 2, 3, 4)[3] describing the development of the techniques, reporting on laboratory engine tests which aided in establishing valid limits in the interpretation of the data, and citing engine histories giving credence to the use of the method as a practical engine-maintenance tool. Currently the emission spectrographic examination of periodic samples of used crankcase oils is being applied by many of the Class I railroads as a primary guide in the continual surveillance of day-to-day operations of many of their diesel locomotives. These railroads often consider this as their most practical means of detecting excessive engine wear, improper air-filter maintenance, water leaks, and distressed bearings.

As is the case with most types of used-oil analyses, the emission spectrographic examination of used oils is most effective and reliable when applied to engines operating under relatively constant conditions. Pipeline-engine operations with their more nearly constant speeds and loads, uniform engine temperatures, infrequent changes in fuel source, and good oil and air-filtration equipment, all contribute to make the oil analyses and data interpretation even more accurate and dependable than in the case of railroad operations where the technique already has proved highly successful.

The Value of the Technique

The technique of utilizing data obtained by the emission spectrographic analysis of periodic samples of used oil from individual engines has value both as a general maintenance tool and as a special technique which provides much needed information in times of engine operating difficulty or when field service tests are in progress. To illustrate in part the versatility of the technique, in the authors' experience it has been used as a surveillance tool to follow engine performance under such circumstances as:

1 In times of known operational difficulty.

2 When redesigned engine components were first put into service.

3 When new oils were being tested for approval.

4 When lubricants were being used under more severe service compared with previously accumulated operational experience.

[1] Phillips Petroleum Company.

[2] Mechanical Engineer, Phillips Petroleum Company. Mem. ASME.

Contributed by the Transportation Committee of the Petroleum Division and presented at the Petroleum-Mechanical Engineering Conference, Dallas, Texas, September 23–26, 1956, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, September 5, 1956. Paper No. 56—PET-30.

[3] Numbers in parentheses refer to the Bibliography at the end of the paper.

The value of the emission spectrographic analysis of used oils as an aid to control of engine operations is based on the fact that the crankcase oils contain trace quantities of various metallic and nonmetallic elements commonly referred to as key wear metals. These elements are present in equilibrium but trace concentrations in the used crankcase oils of engines operating in a normal and satisfactory manner. Changes in concentration levels of one or more of the key wear metals can be related to certain events transpiring in the engine during trouble periods. The reliability of the emission spectrograph as an analytical tool capable of detecting trace amounts of various elements in a material makes it particularly applicable to the rapid detection of the key wear metals in oils since they usually occur in concentrations of a very few parts per million.

This method of used-oil analysis provides information related to events occurring in a particular engine at a particular time. Generally speaking, it is not a technique which can be applied advantageously to only a few engines of a large group in anticipation that the general course of events for all engines can be predicted on the basis of the analyses of samples taken from the small number of selected engines. Used oils, taken from engines in the same service but of different makes, often differ markedly with respect to their equilibrium concentrations of the key wear metals indicative of normal engine operation. Likewise, differences also may arise in the equilibrium concentrations of key wear metals between engines of the same make operated in like service. For example, one engine may consistently show an iron content in the range of 15 ppm; another engine of the same make, 5 ppm; and a third engine of different make operating in the same service may show an equilibrium iron content of 25 ppm, yet all be indicative of satisfactory engine operation. The fact that such differences in key wear-metals content exist between engines does not impair the effectiveness of this used-oil-analysis method when the engines are surveyed on an individual basis.

The Mechanics of Applying the Technique

The practical considerations in obtaining and handling used-oil samples intended for emission spectrographic analysis are not unlike those followed in obtaining acceptable samples for other forms of used-oil examination. As is always the case, the samples analyzed should be representative of the oil in the engine crankcase and should be handled as expeditiously as practical. Whenever possible, it is recommended that the used-oil samples be taken from the circulating oil stream while the engine is running normally. One of the preferred points for obtaining samples is from the oil line immediately ahead of the strainer and filter. The size of the sample taken is not important as long as a representative sample is obtained; only a few grams of oil are actually required for the analysis.

The frequency of sampling is dependent on the degree of reliance to be placed on the data obtained. In the authors' experience with pipeline engines, oil samples have been taken at scheduled intervals of 250 to 1000 engine hr, with the more frequent sampling being followed in special tests and the less frequent sampling in routine surveillance. In addition to the regularly scheduled sampling, when the urgency of the situation demanded it, recheck samples were obtained and analyzed in order that certain trends might be observed with greater frequency. Samples also were taken approximately 200 hr in advance of each anticipated oil drain in order that a recheck sample might be obtained immediately prior to the oil change if the regular sample analysis indicated the desirability of immediate re-examination of the crankcase oil. A new sample for repeat analysis is requested whenever an abnormal concentration of one or more of the key wear metals is indicated unless the deviation from equilibrium concentrations is felt to be so unusual that it might be unsafe to defer an engine-inspection recommendation while awaiting the analysis of a new sample.

To a considerable degree the value of this technique as an aid to preventive maintenance is dependent on how quickly abnormal trends can be reported. This requires that the elapsed time between the date of taking the sample and that of having it analyzed be kept to a minimum. Every effort is made to complete the analysis within a week after the oil is taken from the engine. Expeditious handling of the samples not only will aid in providing data at the earliest possible time, but also will result in more valid data since the possibility of separation of the suspended materials in the oil on standing will be lessened.

It is not the purpose of this paper to relate in detail the analytical procedures employed in determining the concentrations of the key wear metals. The details will vary from laboratory to laboratory depending upon the type of emission spectrographic equipment available. To the spectroscopist, the emission spectrographic procedure to be used is a relatively simple one, especially in view of the fact that the number and frequency of samples lends itself to setting up routine procedures for their preparation and handling. The technique employed by the authors' company was adapted from that described by R. F. Meeker and R. C. Pomatti (5), modified as necessary for adaptation to existing apparatus and in keeping with efforts to obtain the specific information desired with the least amount of analytical effort. Depending on the number of samples handled at a time, one to two man-hours of laboratory analytical effort are required for the preparation and examination of each sample.

The data obtained from the analysis of the samples are evaluated by maintaining a continuing plot of the concentration of each of the elements versus engine hours for the individual units. Thus, taking into consideration the time of each crankcase-oil change and certain other factors not attendant to normal day-to-day operations, it is possible to ascertain at a glance whether or not the particular engine in question is continuing to operate satisfactorily as evidenced by no more than minor deviations from the nominal equilibrium concentration of each of the wear metals. If trends of increasing concentration of one or more of the key wear metals are evident, their magnitude determines the next action to be taken. When the deviations are consistent so as to show continuing trends of slowly increasing magnitude, it is usually the policy to revert to a more frequent oil-analysis interval and follow subsequent data trends very closely. However, in instances where the deviations have persisted over a long period of time, or because of a very rapid increase over a shorter period, engine inspections are recommended at an early date.

There are many metallic and nonmetallic elements which can be detected by this method. However, the number of these elements which are normally usefully related to engine malfunctioning is not great. In pipeline-engine surveillance the amounts of iron, lead, aluminum, chromium, and silicon are determined. In other operations, it might be desirable to determine the content of other metals, including copper, tin, silver, and boron, to name a few. Obviously, there is little reason to expend effort determining metals contents which offer no information concerning the engine operation. On the other hand, in many instances data on less important metals, but ones which may be related to some engine components, can be determined with little extra effort. Analyzing for a larger number of elements usually is compromised against the time required for the additional determination.

Interpretation of the Data

In the final analysis, the manner of relating used-oil data to engine operations must be based on experience—the experience of having determined for a particular operation what magnitude of deviation from equilibrium concentrations of the various wear

metals is significant in terms of indicating the need for engine inspections. While the absolute values of trace metal concentration may vary considerably from company to company, this does not preclude the citing of certain relationships which exist between increases in metal content and abnormal events occurring within the engine. These commonly noticed relationships are given in the following as a guide to interpretation of the data.

Increasing trends of two or more elements often give important clues in the isolation of areas of malfunctioning within the engine. In most instances, there are several sources of a particular metal in an engine. For example, an increasing iron content may result from wear of piston rings, cylinder walls, crankshaft journals, gears, or other ferrous metal parts. High chromium may result from ring or linear wear, or from water leaks. High copper may come from main or connecting-rod bearings, from thrust washers, or from other alloy bearings in the engine.

Considering chromium content first, serious water leaks are normally indicated by very rapid rises in chromium content, often to values of 100–200 ppm or greater. High chromium content subsequently followed by increasing iron content usually suggests a persistent water leak which is affecting ring and cylinder wear. Less rapid but continuing increases in chromium and iron suggest abnormal piston-ring and cylinder-wall wear which might result from broken piston rings or improper air-filter maintenance.

Higher-than-normal concentrations of lead in the used oil suggest abnormal wear, corrosion, or other modes of removal of the lead from copper-lead bearings. Persistence of a trend of higher lead in successive samples, accompanied by rising copper, points to the probability of an impending bearing failure. High copper without high lead suggests deterioration of bronze thrust washers or bearings.

High silicon content suggests airborne dust contamination. High aluminum may indicate excessive aluminum-piston wear, particularly if accompanied by increasing iron or silicon content. Similarly, abnormal concentrations of tin or silver are indicative of abnormal wear of engine parts containing these metals. In engines where borate-treated water is used as the coolant, boron becomes the indicator of water leaks rather than chromium as previously indicated.

Before leaving this discussion of the interpretation of data, it should be pointed out that it is not the intention of this paper to imply that the emission spectrographic analysis of used oils is a panacea for the detection of all of the operating problems which confront the pipeline-engine operator. For instance its usefulness in the surveillance of babbitt-bearing performance is not known. It appears to be of little, if any, value in detecting impending fatigue failures. This is not unexpected, however, since the technique serves to draw attention to abnormal wear or corrosion conditions, which in most cases are not directly associated with fatigue failures.

Likewise, it should be recognized that consideration must be given certain factors that have the temporary effect of disturbing the equilibrium concentrations of wear metals but should not cause alarm as being indicative of damaging engine wear. It is obvious that a change of crankcase-engine oil will disturb temporarily the equilibrium concentrations of wear metals. To avoid misinterpretation of data, spectrographic analyses are not obtained on samples taken less than 250 engine hr after an oil change. Correction of an abnormal wear condition in an engine sometimes results in a carry-over effect causing continued indication of higher-than-normal metals concentrations for some period of time. In most of these cases, however, the trend is downward and easily recognized. Another factor which can cause higher-than-normal levels of key-metals content is the break-in or run-in wear associated with the operation of a new engine, or of newly replaced engine components. Here the higher concentrations are merely indicative that the desired run-in wear is taking place and the eventual stabilization of the key wear-metals content indicates that the break-in process has been completed.

Examples of Typical Data and Service Experience

In Figs. 1 through 6 examples of typical data trends are shown. Figs. 1 through 4 are taken from experience during the past 2 years while applying the emission-spectrographic-analysis technique to pipeline-engine operations. Figs. 5 and 6 are taken from previous experience with railroad diesel-locomotive-engine operations and are included to indicate the usefulness of the analysis in two areas where, in the authors' experience, similar examples have not yet been encountered in pipeline-engine surveillance.

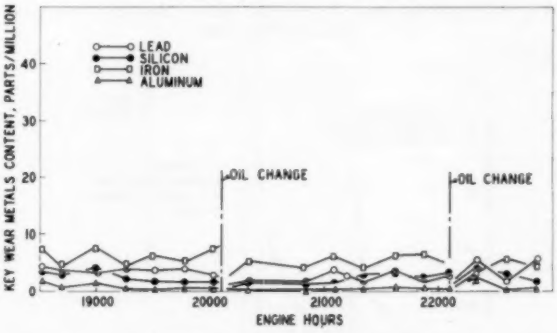


Fig. 1 Typical Key Wear-Metal, Equilibrium-Concentration Curves for a Particular Pipeline Engine Operating Satisfactorily

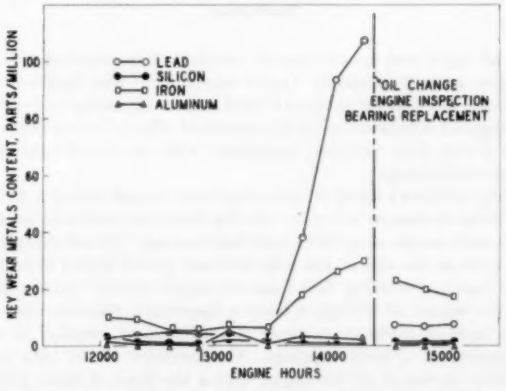


Fig. 2 Example of Bearing Distress as Indicated by Rapidly Increasing Lead and Iron Content

Fig. 1 shows emission-spectrographic-analysis data obtained on periodic samples of crankcase oil taken over a period of several thousand hours of engine operation spanning more than two complete oil-drain periods. In the case shown, engine operation during this period was entirely satisfactory and normal, and is so indicated by the equilibrium concentrations of the key wear metals in the crankcase oil during this period. In the interest of clarity, the concentration line for chromium was omitted. Its concentration line in this case was virtually coincident with that of the aluminum.

Fig. 2 illustrates a very rapid rise in lead content, accompanied by a similar trend of lesser magnitude for iron. The load halves

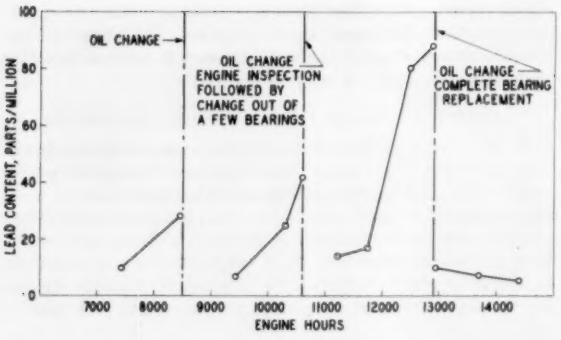


FIG. 3 PROGRESSIVELY INCREASING LEAD CONTENT INDICATING BEARING DISTRESS

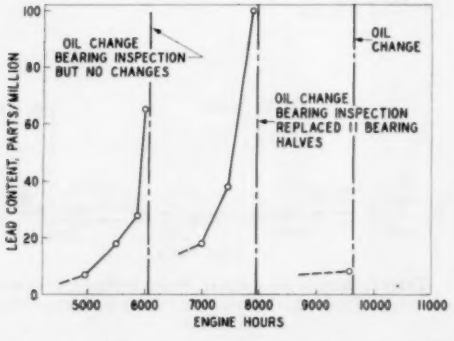


FIG. 4 EXAMPLE INDICATING THE NEED FOR A SECOND BEARING INSPECTION

of all main and connecting-rod bearings were replaced in this engine at approximately 14,400 engine hr. The higher-than-normal, but decreasing trend in lead content following the bearing changeout is believed to be the combined effect of carry-over and the run-in wear normally associated with the initial operation with new bearings.

Fig. 3 shows a trend of increasing lead content during a period of three oil-change intervals. In this instance, concentrations of key wear metals other than lead were normal. A lead content of 28 ppm at the end of the first oil-drain period shown indicated the need for watching data from the engine closely. At the time of the second oil change, a bearing inspection, recommended on the basis of continual increase of lead content, resulted in the replacement of some bearings. Continuation of the high lead content in the oil of this engine during the third oil-drain period indicated the need for further bearing replacements, and a subsequent inspection of the engine at approximately 12,900 engine hr resulted in a complete bearing changeout. This final corrective action resulted in the lead content dropping almost immediately to a normal low value.

Fig. 4 indicates the sensitivity of the method in detecting impending difficulties in their early stages and also illustrates a special value of this sort of engine surveillance—that of suggesting the need for an engine inspection at a time when otherwise an engine inspection would not have been contemplated. In this case, a rapidly rising lead content was indicated during the first oil-drain period shown. Because of the steepness of the lead-concentration curve, an immediate bearing inspection was recommended. Personnel inspecting the engine reported that none of the bearings examined was in such condition as to warrant replacement.

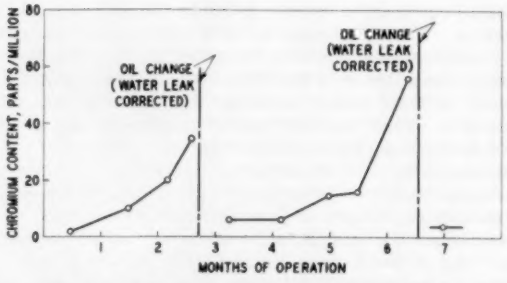


FIG. 5 CHROMIUM-CONTENT TRENDS INDICATING WATER LEAKS DURING TWO SUCCESSIVE OIL-DRAIN PERIODS FOR A LOCOMOTIVE DIESEL ENGINE

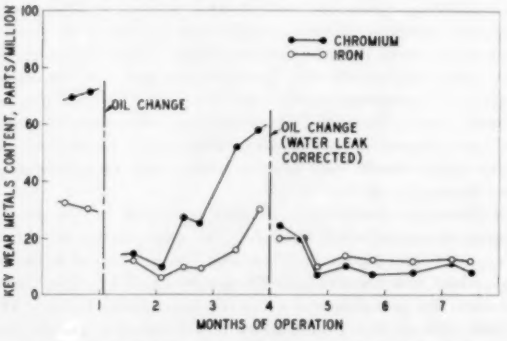


FIG. 6 EXAMPLE OF EFFECT OF PERSISTENT WATER LEAK INDICATED BY RISING CHROMIUM CONTENT ACCOMPANIED BY INCREASED IRON

Ordinarily then, several thousand hours would have elapsed before another bearing inspection would have been made. However, the concentration of lead in the subsequent samples continued to indicate an alarming lead content. Insistence on a second bearing inspection, substantiated by the emission spectrographic data, resulted in the replacement of 11 bearing halves. From that time on, engine operation with respect to bearing condition has been entirely satisfactory. Without the key wear-metal analysis, it is possible that operation of the engine with the defective bearings would have continued until such time as a loss of oil pressure or pounding of the bearings was noticed, possibly incurring damage beyond that which could be corrected by a simple bearing changeout.

Fig. 5 illustrates the manner in which water leaks are indicated by an increasing trend in chromium content. In this instance, the increasing chromium content of oil samples taken from a diesel-locomotive engine is shown. Corrective action and an oil change resulted in satisfactory operation for a period of time, but was followed by another period of water-leak difficulty. These data do not indicate a severe water leak since concentrations of chromium in excess of 100 to 200 ppm are often encountered almost instantaneously when serious water leaks occur.

Fig. 6 illustrates a case of a continuing water leak as indicated by high chromium and iron contents at the time of an oil drain, followed by increasing trends for these metals during the subsequent period of operation. Corrective action at the time of the second oil-change period shown resulted in a return to normal for the concentrations of chromium and iron, indicating that the water leak had been eliminated.

CONCLUSION

In the authors' experience, the technique of utilizing the emis-

sion spectrographic analysis of used oils has been a most useful tool in the surveillance of pipeline-engine operations, particularly during periods when copper-lead bearing performance was in question. The intrinsic value of this type of used-oil analysis lies in the fact that the occurrence of abnormal events in the engine are indicated soon after engine distress or malfunctioning is first encountered, thus suggesting the need for corrective action at a time usually far in advance of probable engine failure.

Acknowledgments

The authors take this opportunity to acknowledge the fine work of R. Q. Gregg, W. D. Purvis, and others of the Research Division without whose analytical assistance this program of emission spectrographic examination of used oils could not have been carried out.

BIBLIOGRAPHY

1 "How Spectrographic Analysis Controls Diesel Engine Maintenance," by H. R. Sennstrom, *Railway Mechanical and Electrical Engineer*, vol. 126, April, 1952, pp. 65–68.

2 "Symposium on Spectrographic Analysis of Diesel Engine Lubricating Oil," a booklet, American Locomotive Company (now ALCO Products, Inc.), Schenectady, N. Y., 1952, pp. 1–76.

3 "Symposium on Diesel Locomotive Engine Maintenance," a booklet, American Locomotive Company (now ALCO Products, Inc.), Schenectady, N. Y., 1953, pp. 1–33.

4 "Rates of Wear in Railroad Diesel Engines," by Ray McBrian and L. C. Atchison, paper presented at SAE National Diesel Engine Meeting, St. Louis, Mo., November 1, 1949. For condensation see "Seven Causes of Railroad Diesel Engine Wear," *SAE Journal*, vol. 58, July, 1950, pp. 18–21.

5 "Spectrographic Determination of Trace Elements in Lubricating Oils," by R. F. Meeker and R. C. Pomatti, *Analytical Chemistry*, vol. 25, January, 1953, pp. 151–154.

Materials Aspect of Some High-Temperature, Refinery-Piping Applications

By E. A. STICHA,[1] CHICAGO, ILL.

New developments in refinery processing have posed many problems for materials engineers through the years. The refining industry was one of the first to require metals for carrying out large-scale operations at high temperatures. It is the aim of this presentation to outline in a general way the materials most commonly used in these applications and to discuss more specifically some of the problems which have been encountered in the past ten years.

ANY consideration of materials for refinery processes must of necessity be restricted in its scope since these operations involve such a wide variety of fluids under diverse conditions of temperature, pressure, corrosion, erosion, and the like. This has been especially true with the entrance of many oil companies into the petrochemical field. Thus, for purposes of this paper, corrosive effects which are often all-important in selecting materials of construction will be accorded incidental attention only. Rather, the emphasis will be on the mechanical and structural properties of the more common metals in use today since, to a large extent, these characteristics reflect ability of materials to retain fluids under extreme pressure and temperature conditions.

Materials for high-temperature service are selected primarily on the basis of their strength (short-time, creep, and rupture) although it should be evident that other characteristics such as ductility, corrosion and erosion resistance, structural stability, weldability, and so on, must be adequate. It is fortunate indeed that many of these requirements are fulfilled in an economical manner by one basic material, namely, carbon steel. We are prone to overlook the versatility of carbon steel as offered by variations in method of manufacture, composition, and heat-treatment. So adaptable is this simple iron-carbon alloy that in some instances it is employed for temperatures up to 1100 F though admittedly at low pressure and in reducing atmospheres.

Where greater fluid pressure must be contained in this temperature range, a higher-strength material is available in the form of carbon-molybdenum steel. This low-alloy steel has essentially the same desirable characteristics as carbon steel in addition to the improved high-temperature strength. Its field of greatest application is in the temperature range from 700 to 1100 F.

Processing conditions often require greater high-temperature corrosion resistance than is obtained with the two aforementioned steels and, to meet this need, a wide variety of chromium-molybdenum alloy steels is available. Use of these materials may start at about 800 F and extend to 1100 F in the case of $^1/_2$ and 1 per cent chromium steels (with $^1/_2$ per cent molybdenum) or to about 1300 F with 9 Cr, 1 Mo steel. Included in this general category is the well-known 5 Cr, $^1/_2$ Mo steel which has served the industry so admirably for approximately 30 years.

In some cases, straight chromium steels of greater alloy content have been employed to cope with particularly corrosive fluids. These materials with chromium levels of about 12, 17, and 27 per cent are not ordinarily used for high-pressure applications above 1000 F owing to lack of strength but may be utilized for even higher temperatures because of their superior oxidation resistance. The 12 and 17 per cent chromium steels, falling within the stainless-steel classification, have been used extensively for valve trim.

For extreme-pressure conditions, occasionally to temperatures as high as 1500 F, engineers often have resorted to use of the austenitic chromium-nickel steels because of their high creep and rupture resistance. These steels contain 18 per cent or more chromium with 8 per cent or more nickel and may be modified by additions of other elements to confer special characteristics. In view of the excellent high-temperature corrosion resistance of the austenitic steels and of the straight chromium steels, it is not surprising that these alloys are sometimes used as lining materials for less costly steels.

Temperature ranges designated for the various steels are subject to modification as conditions and judgment may warrant. The materials mentioned are not, of course, the only ones applied in high-temperature refinery operations but are believed to be the more important alloys in use today. Attention will now be directed to some of the more recent developments in connection with the use of these steels.

Graphitization

In 1943, a welded carbon-molybdenum steel main-steam line fractured next to a weld. This line had been in service carrying 925 F steam for about five years and examination showed that structural breakdown of the carbide phase in the steel, resulting in segregated graphite formation at the edge of the weld heat-affected zone, led to the failure. While this incident caused great concern among steam-power engineers, the petroleum industry with similar material operating under comparable temperature conditions seemed little perturbed. However, several years later when investigation of a cracked reactor revealed random graphitization, an extensive survey of high-temperature equipment was undertaken, results of which have been reported recently by Wilson.[2]

Both carbon and carbon-molybdenum steels were found to be susceptible to development of graphite which occurred usually in a random pattern throughout the steel but sometimes as a heavily segregated formation paralleling the weld-fusion line. One of the prime variables responsible for the carbide instability was reported to be steel-making practice and, in this connection, a large aluminum addition for deoxidation was found to promote graphitization. It also was noted that higher operating temperatures favored graphite formation. A few instances of seemingly anomalous behavior were observed. This experience is essentially the same as that which came out of a study of steam-power piping materials.

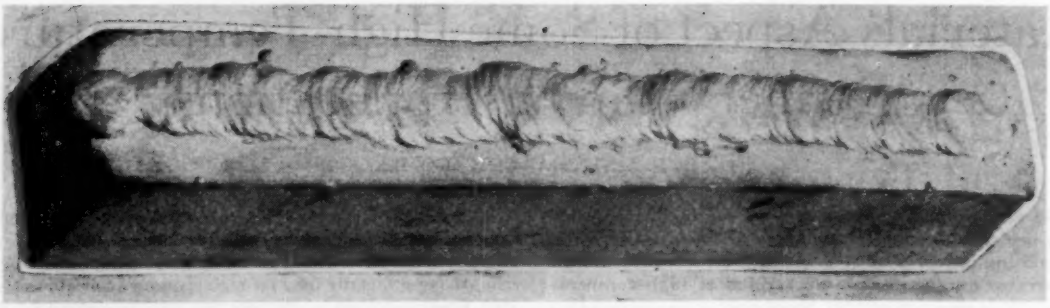
---

[1] Supervising Engineer, Engineering Laboratories, Crane Company.

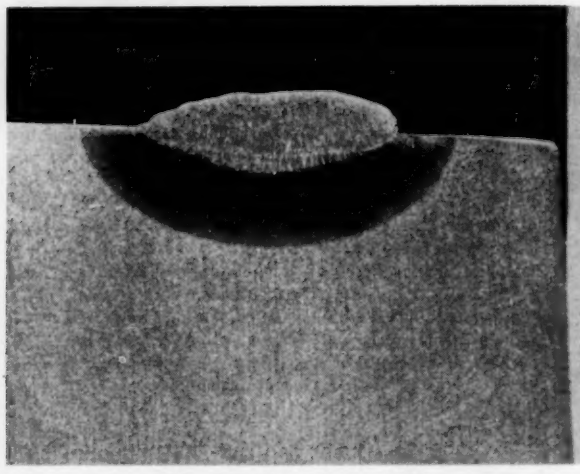
Contributed by the Materials Committee of the Petroleum Division and presented at the Petroleum-Mechanical Engineering Conference, Dallas, Tex., September 23–26, 1956, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, July 5, 1956. Paper No. 56—PET-21.

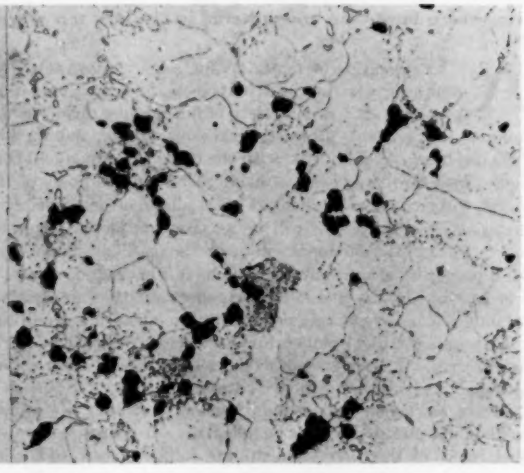
[2] "Graphitization of Steel in Petroleum Refining Equipment—Summary Report of the Subgroup on Deterioration of Steels at High Temperature to the Subcommittee on Unfired Pressure Vessels," by J. G. Wilson, Proceedings of the API, Section III—Refining, vol. 35, 1955, p. 209.



(*a*) Typical weld-bead specimen for graphitization studies



(*b*) Cross section of weld-bead specimen (×4)



(*c*) Graphitization of specimen after aging (×500)

FIG. 1 LABORATORY TESTING OF GRAPHITIZATION SUSCEPTIBILITY

Considerable laboratory effort has been expended in studying the graphitization phenomenon. In general, this work has followed along three channels; namely, evaluation of variations in melting practice, heat-treatment, and composition. The first two approaches have not been particularly fruitful in suggesting practical means of combatting graphite formation in service and the third, the use of alloying elements, has been generally adopted by industry.

The usual procedure in laboratory studies is to deposit a weld bead along a chilled bar of steel, expose the sample at a temperature known to accelerate development of graphite, and then examine the structure periodically under a microscope, Fig. 1. Miniature bend and/or tensile tests may be conducted to assist in the evaluation of materials. Results of these tests are in good agreement with service experience except that the laboratory tests rarely, if ever, develop an intergranular-graphite segregation. A summary of some of the work undertaken at the Crane Company Engineering Laboratories is shown in Fig. 2. Here the degree of conversion of carbide to graphite has been estimated after various exposure times.

Within the scope of this investigation, it is immediately apparent that there is relatively little difference in the behavior of steels made by various melting processes except in the case of bessemer steel. This material exhibits unusually good resistance to graphitization but unfortunately suffers from certain other defects which render it unsuitable for use in critical piping installations. The effect of aluminum additions is also quite pronounced. One might infer from these test results that absence of aluminum will guarantee immunity from graphite formation, but even those steels which are deoxidized with silicon only develop graphite after more prolonged aging. Post-weld heat-treatment is shown to exert a minor influence on graphitization susceptibility. Some treatments have been shown to retard graphite formation but no practical procedure eliminates it.

The role of several alloying elements in carbon-molybdenum steel is shown in Fig. 2 also. Increasing molybdenum content from the nominal $^1/_2$ per cent level to 1 per cent imparts little added resistance to graphitization. The same applies to addition of nickel to the extent of $1^1/_4$ per cent and, though not shown, vanadium is of little help in combatting graphitization. Only chromium and manganese of the more common alloying elements are effective in retarding graphite formation and, of these, chromium is the better. Thus it is that chromium-molybdenum steels are now being specified for high-temperature piping where previously carbon-molybdenum steel was employed. On the basis of equipment surveys and laboratory tests, it is general practice to restrict welded carbon-steel piping to a maximum of 750 F and carbon-molybdenum steel to 850 F. Another consequence of the graphitization problem with these steels and the resultant limited range of application of carbon-molybdenum steel is the apparent elimination of cast valves of this material as a regular stock item.

The effect of residual amounts of certain elements in steel may

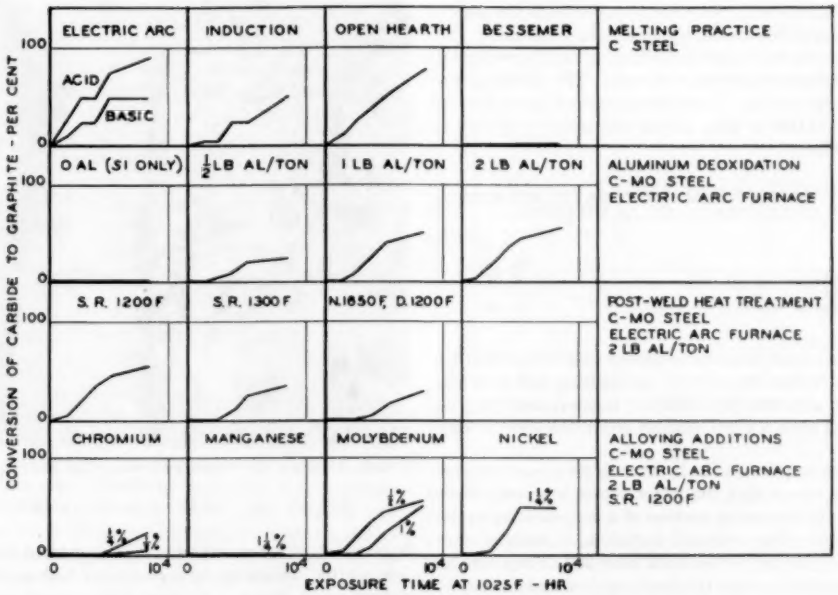


FIG. 2 GRAPHITE FORMATION AS AFFECTED BY VARIOUS CONDITIONS

account for the anomalous behavior of some samples with respect to graphitizing tendency. However, it seems more than likely that elements which are not usually determined analytically may be responsible for this condition. The remarkable resistance to graphitization of bessemer steel also suggests this possibility, and nitrogen, which runs much higher in bessemer than in other steels, has been shown by several investigators[3, 4] to inhibit graphite formation. The possible influence of nitrogen may be modified by the presence of strong nitride formers like aluminum, chromium, and so on.

LIMITATION OF VARIOUS CARBON STEELS

It is sometimes a source of confusion to engineers why certain carbon steels are assigned allowable stress values in codes to

[3] "Effect of Certain Elements on the Graphitization of Steel," by R. J. Fiorentino, A. M. Hall, and J. H. Jackson, Trans. ASME, vol. 76, 1954, p. 1123.

[4] "Roles of Aluminum and Nitrogen in Graphitization," by E. J. Dulis and G. V. Smith, Trans. ASM, vol. 46, 1954, p. 1318.

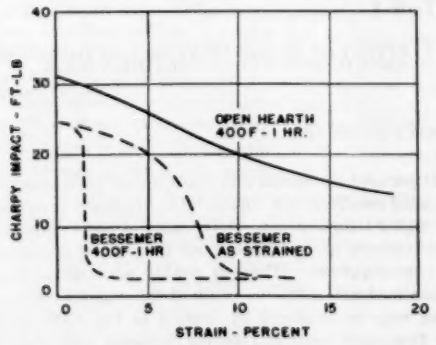


FIG. 3 STRAIN-AGING BEHAVIOR OF BESSEMER AND OPEN-HEARTH STEELS

1100 F (e.g., ASTM A 106 pipe) while others are restricted to only 400 F (e.g., ASTM A 53 bessemer pipe). This situation reflects, to a great extent, differences which result from steelmaking practices. Such characteristics are not revealed by room-temperature tensile tests but may be indicated by a maximum in the curve of tensile strength versus temperature, usually in the region of 400 to 500 F. Perhaps a better way of showing such behavior is by impact tests on specimens previously strained, either at room temperature or in the so-called blue-brittle range. Tests of this sort reflect material behavior at points of stress concentration such as notches, sharp corners, and abrupt changes of section and thus are important in determining applicability of a material. Fig. 3 shows the marked strain-aging behavior of bessemer steel as contrasted to open hearth; the effect is even more pronounced if the straining be performed in the 400 to 500 F range.

CARBON LIMITS FOR LOW-ALLOY CASTING STEELS

An indirect consequence of the graphitization problem with carbon and carbon-molybdenum steels has been a general lowering of permissible carbon content in the alloy casting steels. Prior to 1946, most of the casting steels of ASTM A 157 and ASTM A 217 permitted carbon to a maximum of 0.30 per cent. With the need for chromium additions for combatting graphitization and the desire to maintain welding characteristics, the newly developed alloys were specified usually with 0.20 per cent maximum carbon. Thus casting steels have substantially the same composition and mechanical properties as the tubular products to which they may be joined. Along with this change also comes slightly improved ductility, toughness, and creep strength.

SERVICE EMBRITTLEMENT OF FERRITIC STEELS

While the great majority of carbon and low-alloy steels in high-temperature refinery service retain their desirable properties to a large extent, instances have been reported of embrittlement occurring after extended periods of operation. This condition has been observed in carbon, carbon-molybdenum, and 5 chro-

mium, 1/2 molybdenum steels, and could be expected to develop in other chromium-molybdenum alloy steels also. Generally the affected material shows no change in hardness or tensile properties but the impact resistance is decreased sharply. The results shown in Table 1 for a 5 chromium, 1/2 molybdenum-steel pipe removed from service after 74,000 hr with a fluid temperature of 1020 F demonstrate the effect.

TABLE 1 EFFECT OF HEAT-TREATMENT ON IMPACT RESISTANCE OF EMBRITTLED 5 CR, 1/2 MO STEEL

	Charpy impact, ft-lb
After service	18, 19 1/2
Heated at 1300 F—4 hr	55, 57
Heated at 1100 F—4 hr	34
Normalized 1650 F, drawn 1300 F	65, 62

These values also show that the condition can be removed by heating below the critical range or by normalizing and drawing. It is interesting to note that a substantial improvement may be realized by heating even a small amount above the service temperature.

This phenomenon is quite similar in many of its aspects to temper embrittlement except that the latter occurs in a very short time; i.e., during the tempering portion of a heat-treating cycle. In contrast, a sample of the restored 5 chromium, 1/2 molybdenum steel retained its impact properties even after 262-hr exposure at 950 F. It seems possible that the two types of embrittlement may stem from the same source but this need not be so. Investigators have attributed temper embrittlement to various causes among which may be listed changing carbide solubility with temperature and presence of residual impurity elements such as antimony, arsenic, phosphorus, tin, and titanium. Molybdenum is generally considered to be a specific remedy for temper embrittlement but as much as 1/2 per cent appears inadequate for the service embrittlement described herein. No positive metallographic evidence of an embrittling intergranular phase has been forthcoming in connection with this type of service embrittlement. It is fortunate in view of our lack of knowledge on this subject that the deficiency is not encountered more often.

Welding of Chromium-Molybdenum Steels

Most chromium-molybdenum steel refinery-piping elements are joined by welding now. In the past it was not uncommon to weld some of these joints with austenitic-steel electrodes, usually 25 Cr, 20 Ni alloy, and dispense with stress-relieving. It should be pointed out that this procedure leaves alloys like the 5 chromium, 1/2 molybdenum steel in a hard and relatively notch-sensitive condition in the heat-affected zone adjacent to the weld. While this procedure has been employed successfully in many cases with light-wall tubing, its general application should be considered carefully. Probably trouble was avoided in these instances because of ability of the ductile weld deposit to absorb line strains. In view of developments which have yielded improved procedures and electrodes of compositions matching the base material, there seems to be little need for continued usage of austenitic-steel electrodes for welding of ferritic-steel piping except under unusual circumstances for field work.

For elevated-temperature service above about 800 F, the practice of joining ferritic-steel piping with austenitic-steel welds presents another potential hazard. Where the austenite phase is in contact with ferrite, carbon will diffuse to the higher alloy (producing a carbide-rich zone at the interface) and away from the low alloy (leaving a weak decarburized area adjoining the carbide). The net result is a brittle zone immediately adjoining the fusion line which is subject to failure under unusual loading conditions. This same situation will result in a ferritic-steel system if differences in chromium content exist between the weld and the base metal. Carbon diffusion is even evident in dissimilar metal joints after application of a post-weld heat-treatment above the critical range, Fig. 4. In view of the danger attending an unexpected failure during refinery operations, use of dissimilar metal weldments should be weighed very carefully.

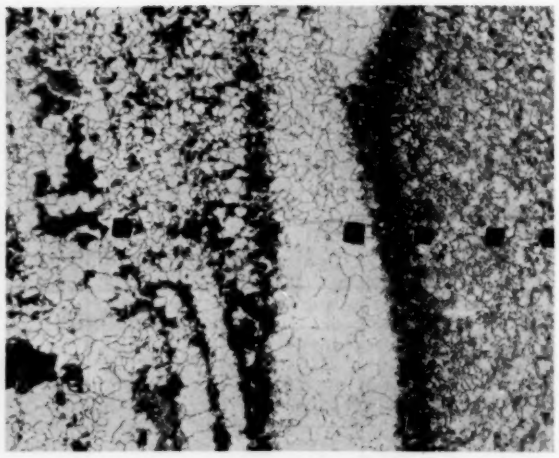


Fig. 4 Carbon Diffusion in Dissimilar Metal Weld Annealed at 1650 F

(Left side, 1 per cent Cr—right side, 5 per cent Cr; ×150.)

Since welding has become such an important part of refinery fabrication and erection, it is well to note that the newer welding techniques are applicable to the chromium-molybdenum steels even at the higher chromium levels. Automatic, submerged-arc welding is employed on a routine basis in shop fabrication. Special attention should be given to use of the inert-gas-shielded-arc process for making the first pass of pipe welds without backing rings and with or without added filler metal or inserts. The process has particular merit where complete elimination of crevices and projections is desired.

885 F Embrittlement in High-Chromium Steels

As noted previously, the high-chromium steels are used to some extent because of their superior corrosion resistance. They are subject in varying degrees to a loss of ductility which has been termed "885 F embrittlement" because its maximum effect is noted in this range. The 27 per cent chromium steel is most susceptible to this form of deterioration. Material properties such as hardness, tensile strength, and ductility undergo changes but the most profound effect is in impact resistance as may be seen in Table 2.

TABLE 2 EFFECT OF EXPOSURE AT 850 F ON IMPACT RESISTANCE OF 27 PER CENT CHROMIUM STEEL

	Charpy impact, ft-lb		
	80 F	200 F	850 F
Original—annealed	4 1/2	33	38
Aged at 850 F for 1000 hr	0	0	6 1/2

The 27 per cent chromium iron tends toward brittleness even in the annealed condition but considerable toughness is available at slightly higher temperatures. After aging at 850 F, however, the material becomes quite brittle, both at room temperature and elevated temperatures. There is neither microscopic nor x-ray diffraction evidence as to the cause of the embrittlement. This condition may be removed by heating in the 1000 to 1200 F range. Prolonged heating, however, between 1000 and 1400 F results in sigma-phase precipitation which, in turn, locally depletes chromium from the matrix and reduces corrosion resistance.

This process may be reversed and toughness together with corrosion resistance restored by heating above 1500 F and cooling rapidly.

No completely satisfactory explanation of 885 F embrittlement is available at this time. Cold-working prior to exposure in the embrittling range enhances the effect and suggests a relationship to sigma-phase formation. Similar changes occur in 17 per cent chromium steel but to a lesser degree. The AISI Type 410 Steel (12 per cent chromium) seems to be entirely free of such embrittlement but the aluminum-bearing modification of this alloy, Type 405, is susceptible to some extent though possibly in a somewhat lower temperature range. Application of the high-chromium steels to elevated-temperature service is understandably restricted because of the embrittling structural changes.

Experience With Austenitic Steels

Apart from usage where only corrosion resistance is desired, the chromium-nickel-iron alloys are employed for extreme-pressure and temperature service.

Although high creep strength and rupture strength are indicated by laboratory tests, these materials suffer from some serious defects. Lack of structural stability in elevated-temperature service as evidenced by intergranular carbide precipitation and sigma-phase formation, both of which reduce corrosion resistance and promote embrittlement, are well known. A relatively stable carbide structure for short exposures to high temperatures, as required for welding applications, is obtainable through the addition of carbide-forming elements like columbium and titanium. A measure of immunity from the deleterious effects of carbide precipitation is available through utilization of what has been termed a stabilizing heat-treatment. In this treatment, carbide is permitted to precipitate and agglomerate at a high enough temperature (1525 to 1650 F) to allow diffusion and homogenization of the chromium content. Thus corrosion resistance is not seriously impaired.

In casting grades, the wider limits on chemical composition make for greater variability of structure and, hence, of properties. Fully austenitic steel is obtained if austenite formers (Ni, C, N, Mn) run high while ferrite formers (Cr, Si, Mo, Cb, Ti) run low. The structure may contain considerable ferrite if the reverse is true and it is from this ferrite that sigma phase forms most readily. The effect of exposure at 1200 F on impact properties of two of the more popular casting grades of austenitic steels is shown in Fig. 5, and a comparison of the loss of toughness in the high and low-nickel modifications of Grade CF8M steel illustrates this point.

Welding of austenitic steels in light sections has not posed any particular problem, especially if stabilized or if extra low-carbon grades are used to avoid carbide precipitation. With heavier wall thicknesses, increased restraint on the weld deposit has led to development of microcracks in the weld and, in some cases, to macrocracks. Fully austenitic-steel weld deposits are most susceptible to formation of microcracks and the condition may be alleviated by balancing the composition of the weld deposit so it contains a small amount of ferrite. A structure of this type is, of course, subject to the same deterioration at elevated temperatures as similar casting steels. In making welds, use of preheat usually is avoided, interpass temperature is limited to 250 F and normally no stress-relieving is practiced. If corrosion resistance or dimensional stability is an important consideration, weldments may be fully annealed or given a stabilizing heat-treatment.

In heavy-wall AISI Type 347 steel steam-power piping, still another difficulty has been encountered in the form of base metal cracking a short distance from the fusion line. Cracking of this

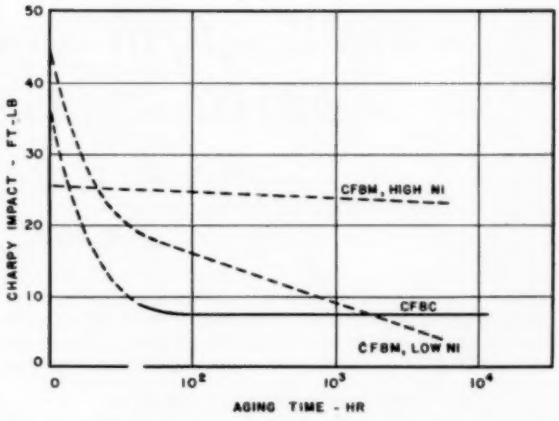


Fig. 5 Effect of Exposure at 1200 F on Impact Resistance of Cast Austenitic Steels

sort is more prone to occur in a coarse-grained material. To date no such trouble has been reported in refinery piping but it behooves engineers to take cognizance of the possibility of such an occurrence.

Failures of cast austenitic-steel valve bodies in thermal-cyclic processes have been reported by Holmberg[5] and attributed to thermal-shock conditions. In this service, the inner surfaces of the heavy-wall bodies are subjected to rapid heating by hot fluid and, because of the poor thermal conductivity and high expansivity of austenitic steel, the material upsets itself. Later, as the temperature equalizes, the inner skin is under a tensile stress of greater magnitude than if the heating had been accomplished slowly and uniformly. Repeated cycling of this sort leads to internal cracking.

This explanation of the failure seems quite plausible although alternate suggestions have been made involving (*a*) stress corrosion by a high-boiling-point condensate or (*b*) a thermal-fatigue process with lowered strength under cycling-temperature conditions. While engineers may not agree as to the mechanism, it has been observed that use of ferritic steel and lessening the temperature swings give longer valve life.

During shutdown periods, austenitic steels are subject to stress-corrosion failures, especially in the presence of chloride ions. It is not always readily apparent where the two essential ingredients of such a process, stress and a specific corrosive agent, come from. Sometimes the high coefficient of expansion of austenitic steels (approximately $1^1/_2$ times that of ferritic steels) is the source of stress while, in other instances, residual stresses in weldments may be responsible. Corrosive agents may come from brackish water used for hydrostatic-pressure testing, cleaning reagents, or condensates. Applications involving ferritic steels clad with austenitic alloy warrant examination of the possibility of developing stress-corrosion-type failures.

The materials discussed in this paper do not, of course, embrace all the metals that are utilized in high-temperature refinery piping. It is apparent that there is a real need for improved materials for a number of the operations. However, engineers in the petroleum industry are a most progressive group and we can be assured that full advantage will be taken of new metallurgical developments.

[5] "Experience With Austenitic Steels in High-Temperature Service in Petroleum Industry," by M. E. Holmberg, Trans. ASME, vol. 73, 1951, p. 733.

Operating Considerations in Application of Gas-Turbine-Driven, Centrifugal, Pipeline Compressors

By A. L. VAUGHAN,[1] OMAHA, NEBR.

Northern Natural Gas Company has operated two gas-turbine stations through three winter heating seasons. Each of these stations consists of four 5700-hp gas-fired turbine-driven, centrifugal compressors arranged in series. The intent of this paper is to present the fundamental reasons behind the decision to use turbines in this application and to reflect some light on anticipated related operating and maintenance costs. Sufficient operating time has transpired to arrive at some definite conclusions as to operating costs and the reliability of the machines. However, there is insufficient experience to predict future maintenance costs accurately. Included herein are results of the recent major inspection and overhaul of two gas turbines.

THE company with which the author is associated celebrated its 25th Anniversary in 1955 as a natural-gas-transmission company. During this quarter of a century, the development of the natural-gas industry has made possible the economic use of a product which, in many instances, had been a waste product of the petroleum industry. Since the initiation of large-diameter interstate pipeline systems, the growth of the natural-gas industry has been phenomenal. Our own company is in many ways typical of the industry. Our transmission-system capacity has doubled in the past five years and quadrupled in the past ten years. These expansion programs often have made necessary a new design for increased capacity before the previous transmission-system design could be constructed and the effectiveness actually determined.

This rapid expansion together with increasing costs of compressor-station facilities has left the industry in a position where it must be ever cognizant of new developments in equipment and techniques. In many cases the industry has been obliged to accept a new development on a large scale long before it had been thoroughly proved and accepted. Such was our company's position when the decision was made to install gas-fired, turbine-driven, centrifugal compressors in the 1953 expansion program. This decision involved the purchase of over 34,000 hp to be installed in two new stations to be located between existing reciprocating stations. To our knowledge, there were only two experimental gas-fired, turbine-driven centrifugal compressors in operation prior to initial placement of our order for six turbines in 1952. However, there was one large-scale turbine-expansion program then under construction on another company's transmission system. The decision to install turbines in very critical points in our transmission system was not one that could be made entirely on the basis of industry experience with the application. We were convinced that the basic design and principles were good and time must tell the rest.

[1] Superintendent, Compressor Stations and Processing Plants, Northern Natural Gas Company.

Contributed by the Refining Committee of the Petroleum Division and presented at the Petroleum-Mechanical Engineering Conference, Dallas, Tex., September 23–26, 1956, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, June 28, 1956. Paper No. 56—PET-13.

SYSTEM-DESIGN CONSIDERATIONS

This paper is prepared primarily for the purpose of giving a résumé of operating experience with gas-fired-turbine prime movers driving centrifugal gas compressors. To evaluate this operating experience, it is essential to consider first some of the characteristics peculiar to our transmission system and its operation along with some of the considerations leading to the selection of turbines and centrifugal compressors.

In our opinion, there were three basic methods of increasing capacity which were applicable to our transmission system:

1 Additional compressor horsepower at presently existing stations.

2 Increase existing pipeline volumetric capacity to decrease the interstation pressure drop.

3 Install additional compressor stations (usually at mid-points) between presently existing stations.

In actual practice it is seldom feasible to increase the capacity of an entire transmission system by any one of these three methods alone. Such was the case in the 1953 expansion program when our company increased its salable capacity from 675 to 1007 million cu ft of gas per day. Engineering design in final form required 69,000 additional horsepower and 301 miles of parallel main-line piping in addition to the two new stations called Macksville and Tescott (Kansas). The stations were to be located at mid-points between existing reciprocating installations. Choice of locations was dictated by a requirement for a better horsepower pipeline balance.

The next step was to select a prime mover-compressor combination which would best fulfill the requirements of the current design and would be adaptable to needs in the foreseeable future. Initially, the selected equipment must compress 835 million cu ft of natural gas per day from 518 to 745 psig, a compression ratio of about 1.4. Also, the equipment must be adaptable to handle future volumes in excess of one billion cubic feet per day. It was desirable to provide economy of operating labor cost in anticipation of a seasonal fluctuating load factor. Finally, the installed cost was not to be excessive as compared with any other type. The equipment combinations considered are as follows:

1 Gas engines and reciprocating compressors.
2 Gas engines and centrifugal compressors.
3 Steam turbines and centrifugal compressors.
4 Gas-fired turbines and centrifugal compressors.

The combination of gas-fired turbines and centrifugal compressors was selected for the following reasons:

Centrifugal Compressors. Centrifugals were preferred to reciprocating compressors for operations at very low ratios. The pressure drop across the suction and discharge valves of a recipro-

cating compressor is relatively unimportant at higher ratios; however, at low ratios these losses become significant in compressor efficiency. Inlet and outlet losses of a centrifugal compressor are negligible. There is, therefore, a saving in brake horsepower required to compress 1,000,000 cu ft of gas. The large throughput volumes also were factors in favor of centrifugals as the required horsepower could be utilized in a small number of large-horsepower machines—in our case, three compressors.

Ambient Temperature Versus Turbine Horsepower. A fundamental characteristic of gas-fired turbines is their ability to develop more horsepower at low ambient temperatures. This increase in horsepower conforms very well to our load requirements during winter months as is illustrated in Fig. 1. Our heating is placed under heavy load only during those months when low ambient temperatures make extra horsepower available.

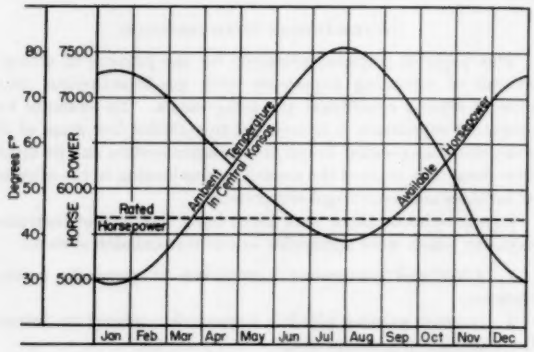


FIG. 1 AVAILABLE TURBINE HORSEPOWER BY SEASON

Personnel Requirements. The turbine stations proposed were to be designed for completely automatic operation. Although there has been much discussion relative to completely unattended stations, it was the company's decision that continuity of service is of paramount importance, therefore any consideration of unattended stations was abandoned. However, economic operation of these plants necessitated the use of as small a staff as possible. It was our opinion that a staff of three operating personnel per tour would meet these service requirements. This staffing is still considerably below what could at that time be anticipated for a reciprocating station. Operating labor is a fixed cost whether the plant is or is not in operation.

Initial Plant Investment. As our company had no past experience with equipment of this nature, it was necessary to make cost studies in much more detail than was our normal practice for preliminary estimates. Completion of these estimates placed the gas turbines in approximately the same relative cost position as the reciprocating plants, which position was later verified by actual cost figures. The three-unit turbine stations were built for $270 per rated brake horsepower while a reciprocating station of similar size was built for $274 per rated brake horsepower.

OPERATIONAL EXPERIENCE

Since our first gas turbine was placed in operation in December, 1953, there has been accumulated in excess of 52,000 operating hours as of March 1, 1956. These hours include two additional turbines installed in the 1954 expansion program. One of these units was installed at each station, making a total of eight turbines and 45,600 installed horsepower.

During the first 27 months, the equipment was available for operation in excess of 90 per cent of the time. Considering that these units were not required for lower load-factor operations, the turbine plants here are compared on a basis of availability rather than on a basis of hours operated. For comparison, similar data have been compiled and summarized in Table 1 for a vertical and a horizontal reciprocating station which also has operated on a similar basis. Although the reciprocating station has a better availability record, it should be noted that the time considered for the turbine includes original station start-up and the early months of operation when the so-called "bugs" had to be worked out of the equipment.

TABLE 1 STATION TIME RECORD

Description	Percent of Total Possible Operating Hours		
	Gas Turbine Station	Horizontal Engine Station	Vertical Engine Station
Hours Operating	33.16	37.78	29.67
* Dispatching Orders	58.35	60.19	68.76
Total Availability	91.51	97.97	98.43
Downtime			
Equipment Failure	4.50	0.30	0.17
Overhaul & Inspection	3.99	1.73	1.40
Total	100.00	100.00	100.00

* Equipment is operable but not required.

Today, we are confident that the over-all reliability of a turbine station will be comparable with our reciprocating stations. We have one station in which all four turbines have operated continuously more than two months without a shutdown. This record might well have been extended had there been continued demand for the equipment.

Operating costs are a prime consideration in the selection of any type of compressor-station installations. Fig. 2 is a summary of operating costs for turbine and reciprocating stations. The stations considered are not a composite of all reciprocating stations on our transmission system but are specific ones which operate on a comparable load factor. These do not represent the costs which we would anticipate in a reciprocating station of today's design with automatic controls and the consequent lower personnel requirements.

Fig. 2 illustrates that a major saving from the installation of turbines has been operating labor. Our turbine stations are currently being operated with three men per tour. An engineer is responsible for operations and is headquartered in the control room which is separated from the turbine room. He is assisted by an auxiliary engineer and an oiler. The auxiliary engineer is

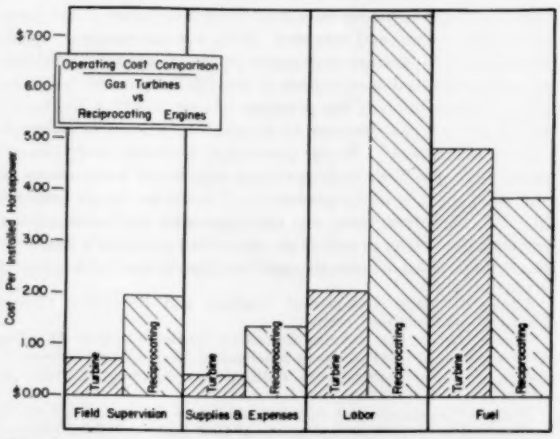


FIG. 2 OPERATING COST COMPARISON—GAS TURBINES VERSUS RECIPROCATING ENGINES

required since our company elected to install separate engine-driven generators for power. He is also responsible for attending heating boiler, air compressors, water-treating equipment, and other appurtenances. The oiler is primarily the engineers' communications contact in the compressor building. He also participates in general maintenance work and other utility service.

A factor favoring turbines is the low oil consumption for that machine as compared with reciprocating engines. The average reciprocating engine on our system develops 5000 hphr for each gallon of lubricating oil while the turbines develop 55,000 hphr per gal. Another factor favoring the turbine plant is the smaller requirement for water make-up and cooling facilities. It should be noted that some of the difference in the amount of cooling involved is not inherent in the two applications, but is because no gas cooling is required at the low compression ratios in the turbine stations.

The fuel costs which are compared in Fig. 2 are total station costs including auxiliaries. The cost figures shown are on the basis of rated horsepower, and do not take into account the extra horsepower developed by the turbine at lower ambient temperatures. With this consideration, the turbine plant is actually in a more favorable position with presently installed reciprocating stations than the graph indicates. Averages throughout the system indicate a station fuel consumption including all auxiliaries as follows:

1 Gas-turbine stations.....10400 Btu/bhphr developed
2 Vertical gas-engine stations..... 9200 Btu/bhphr developed
3 Horizontal gas-engine stations..10600 Btu/bhphr developed

Maintenance

Turbine-plant maintenance requirements and costs are still a relatively unknown factor so far as the company is concerned. Currently one turbine has been inspected and overhauled at each of the two turbine stations. This, however, was not done because of the mechanical condition of the equipment, but we felt it was essential to obtain a basis for anticipating future overhaul requirements. Other considerations favoring this program were to prevent the necessity for overhauling all units the same year and to have trained maintenance personnel when overhaul became a necessity.

Essentially, we feel that the experience gained outweighed the early expenditure. At this time, we were able to use our own operating personnel to conduct the inspection. This afforded a twofold advantage in that we pursued this work without increasing our own staff and provided an opportunity for the operating personnel to improve their knowledge of the fundamental composition of the machine. The only additional help involved was a representative of the turbine manufacturer who was retained throughout the inspection in an advisory capacity. On the basis of the experience obtained at these two locations, we will again this year inspect and overhaul one unit per year at each plant which, if carried out, will make a 4-year overhaul cycle. So long as our fluctuating load factor remains as it is, this cycle will involve approximately 15,000 operating hours per unit between major inspections.

If the two turbines which we have inspected are indicative of what we may expect in the future, the maintenance picture is an encouraging one. The internals of both machines were found to be generally in good condition with some relatively minor exceptions.

Shown in Table 2 is a breakdown of the man-hours expended in major inspection on the No. 1 gas turbine at Tescott. We are convinced that the number of man-hours shown will be less as the personnel become more familiar with inspection and overhaul procedures. The man-hours for major turbine inspection as shown in Table 2 are considerably less on a rated-horsepower basis than those we have experienced at our reciprocating stations.

TABLE 2 MAN-HOURS EXPENDED TO OVERHAUL ONE 5700-HP GAS TURBINE AT TESCOTT, KAN.

Disassembly		
Axial Flow Compressor	46:30	
Turbine Assembly	185:30	
Controls and Accessories	38:45	
Total Man-hours, Disassembly		270:45
Cleaning of Parts		196:15
Assembly		
Axial Flow Compressor	30:30	
Turbine Assembly	289:00	
Controls and Accessories	48:30	
Total Man-hours, Assembly		368:00
Total Man-hours to Overhaul		835:00
Man-hours by Manufacturers' Representative		128:00

When evaluating a maintenance program, a spare-parts stock must be maintained to assure efficient operation. The stock of parts to accomplish this is a very complex thing to define and varies considerably between transmission companies. Northern's spare-parts stock inventory is $5.70 per brake horsepower for turbines which is considerably higher than the $3.15 average for reciprocating stations having vertical gas-compressor engines.

Conclusions

The gas-fired turbines on our system have operated through three heating seasons, in which time the turbines have proved themselves to be equally reliable with reciprocating-compressor stations. The projected operating and maintenance costs are difficult to predict since there has been a variable demand for the turbine facilities. However, the operating and maintenance expenses for the turbine stations were approximately $5.00 per brake horsepower, less fuel, per year as compared to approximately $13.50 for other reciprocating stations which operate on a comparable load factor. The major part of this saving is in operating labor.

From a cost-of-operation standpoint, it is apparent that there is much to be gained by installing facilities with large-horsepower units and to provide sufficient controls and safety devices so that operating labor costs may be held as low as possible. Each installation must be weighed individually to obtain a balance between the cost of operating labor and the need for continuity of service and protection of equipment. We feel that we have satisfied our requirements for this continuity and protection and have still shown a very favorable saving in operating costs.

Our gas-turbine experience has left us convinced that the decision to use gas turbines in this particular application was a good one. Our experience to date is such that gas turbines would be considered in any application in which they could be used. However, recent advances in automatic controls and operation of reciprocating engines have made them very strong contenders in most gas-transmission applications.

The Unsteady Laminar Boundary Layer on a Flat Plate[1]

By SIN-I CHENG[2] AND DAVID ELLIOTT,[3] PRINCETON, N. J.

The problem of a semi-infinite flat plate started initially from rest in an incompressible fluid with an arbitrary velocity without flow reversal is considered in this paper. In reference (1)[4] the unsteady laminar boundary layer on a semi-infinite flat plate started to move with a velocity proportional to t^n, where n is positive, is analyzed in detail. It is shown that a solution, constructed from the leading edge and properly continued into the far downstream region of the leading edge, possesses the proper limiting behavior that either at small t or in the far downstream region the flow field is substantially independent of x and is given by the Rayleigh-type solution. This solution also possesses the proper behavior at very large times. In view of the fact that serious difficulties have been encountered in trying to perturb the Rayleigh-type solution either to obtain solutions in the upstream region or to obtain solutions valid at large times, it is of interest to see whether the method of constructing the unsteady laminar boundary-layer solution as shown in reference (1) can be applied to arbitrary plate velocity and still lead to solutions with proper limiting behavior. This is found to be so if the plate was started initially from rest. In particular, the limiting form of the results for large t produces a solution for the motion of an oscillating plate in a uniform stream. This is of interest at least in so far as the theoretical results can be checked by experiment, not mentioning its connection with a number of practical problems. The effects of compressibility have been considered briefly with Howarth's transformation in reference (2). Unfortunately, the simple energy integrals of Busemann and of Crocco when $Pr = 1$ cannot be extended to unsteady flow with physically sensible thermal condition on the plate. It seems that the temperature field would be better determined numerically for each specific example in question and is omitted from the present paper.

Introduction

PUBLISHED works on the boundary layer with time-dependent relative motion between the plate and air far away from the plate are relatively few. Stewartson's (3) work on the impulsive motion is discussed rather in detail in reference (1). (For more complete references on impulsive motion problems, the reader is referred to reference 2). Lighthill's (4) general linearized treatment is of special interest because it offers a check to the present result after being linearized with respect to the small amplitude of oscillation. Moore's extensive investigations (5) seem to be intended to give a general solution for a wide variety of unsteady laminar boundary-layer problems. His solution was constructed with subtle physical arguments and agrees, at least formally, with the series solution around the leading edge in the present solution which is constructed only for a class of unsteady motions of considerably more restricted nature, and is valid only in a restricted region. The present paper is most concerned with a unified solution for an approximate description of the entire flow field and a discussion of its validity under different circumstances.

It is first noted that in the case of an incompressible fluid, the problem of a flat plate moving with arbitrary velocity $u_w(t)$ into a fluid at rest is equivalent to the one with the plate at rest in an unsteady stream with velocity $u_e(t) = -u_w(t)$. In view of the discussions in reference (1), it is not attempted to construct a solution upstream on the plate by perturbing on the known solution at downstream infinity. Instead, a solution is constructed from the leading edge of the plate and working downstream. As indicated in reference (1), we cannot expect any reduction of the number of independent variables from (x, y, t) to two parameters only. But in order to take advantage of the previous solutions in reference (1), a characteristic series-expansion parameter is tentatively selected to be $\xi = x \Big/ \int_0^t u_e(t)dt$ for the solution around $\xi = 0$. Analysis shows that all physical quantities must be expanded in ascending powers of $\xi_0 = \xi^{1/2}$ with coefficients which are functions of ζ_0 and t. ζ_0 is found to be a generalization of the Blasius parameter, given by $\zeta_0 = y[u_e(t)/\nu x]^{1/2}$. This solution is expected to be valid for sufficiently small values of ξ and is likely to diverge for sufficiently large values of ξ, say $\xi > \xi_i$. In order to establish some criterion of the region in which this series solution may apply, and to determine the flow field with $\xi > \xi_i$, the continuation of the boundary-layer solution into the downstream region from a given velocity profile at a given station is considered. It is found that the viscous-flow field in the vicinity of the plate and naturally the skin friction can be determined by regular perturbation on the known initial velocity profile, near the wall. The determination of the flow field away from the wall is complicated by the presence of a logarithmic singularity which fortunately exerts little influence on the flow field near the plate. This perturbation solution, as in the case of reference (1), enables us to investigate the downstream limiting behavior of the proper continuation of the solution already determined in the upstream region. It is shown that the skin friction approaches the Rayleigh value monotonically as ξ increases. This assures us that no greater error may be involved if we take the Rayleigh value for all the downstream stations with $\xi > \xi_i$. If better accuracy is required, one may either determine more terms in the series expansion around the leading edge to increase the range of validity of the series approximation, or to carry out the perturbation scheme according to the continuation procedure starting from $\xi = \xi_i$. For a number of practical problems such process does not seem to be required.

---

[1] This work is condensed from reference (2) of this report under the same title and is supported by Air Force Office of O.S.R. under Contract No. AF 18(600) 498.

[2] Assistant Professor, Aeronautical Engineering Department, Princeton University.

[3] Research Assistant, Aeronautical Engineering Department, Princeton University; now at N.P.L., England.r

[4] Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Heat Transfer Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS and presented at the Heat Transfer and Fluid Mechanics Institute, Stanford, Calif., June 21–23, 1956.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, May 11, 1956.

The solution is so constructed from the Navier-Stokes equation under the boundary-layer approximation by enforcing the following conditions expressed in terms of the co-ordinates fixed on the leading edge of the plate:

(i) The nonslip condition on the plate surface—the velocity component u parallel to the plate surface and the component v normal to the plate must vanish on the surface $y = 0, x > 0$ for all times $t \geq 0$.

(ii) The velocity component u must be continuous over the entire flow field; in particular, the boundary-layer solution must pass into the uniform incoming stream at all times $t \geq 0$ as $x \to 0$.

In addition, the following conditions must be satisfied by the solution, at least approximately:

(iii) The u-component velocity must pass into $u_e(t)$ at large normal distances away from the plate, $u(x, \infty, t) = u_e(t)$ for all x and $t \geq 0$, and the v-component velocity must be a higher order small quantity.

(iv) The solution must pass into the Rayleigh-type solution at any instant $t \geq 0$ in the flow region very far downstream of the leading edge or at very small $t > 0$ over practically the entire flat plate; i.e., $\xi \gg 1$. The Rayleigh-type solution is the one pertaining to the corresponding problem with a doubly infinite flat plate with no x-dependence under the effect of vorticity diffusion only.

(v) The solution for $u(x, y, t)$ must pass into the steady-state boundary-layer solution when t approaches zero; i.e., the instant when unsteady motion begins to set in. If the plate started from rest, then it must approach a field of no flow.

These conditions are probably as much as one can require with the boundary-layer approximation. With these conditions fulfilled, this solution may be considered as a nice approximate solution of the complete Navier-Stokes equation except possibly for very small x or very small t. The solution obtained in the present paper is applicable only when $u_e(t)$ satisfied the following restrictions:

1 $u_e(0) = 0$ for all x and y to satisfy the condition (v) at $t = 0$.

2 $u_e(t) > 0$ for all $t > 0$ so that the station $\xi = 0$ is always the leading edge of the development of the viscous layer where the upstream flow is uniform.

3 $u_e(t)$ is differentiable indefinitely for all $t > 0$.

Solution Around Leading Edge

Let us take the rectangular Cartesian co-ordinate axes fixed at the leading edge of the plate ($x = 0$) with the plate surface defined by $y = 0, x > 0$. The free-stream velocity $u_e(t)$ is parallel to the plate and is always positive and differentiable at all $t > 0$ and is equal to zero at $t = 0$. A stream function $\psi(x, y, t)$ is introduced so that $u = \psi_y$ is the velocity component parallel to the plate and $v = -\psi_x$, the velocity component normal to the plate. The equation of mass continuity is identically satisfied. The momentum equation according to the boundary-layer approximation becomes

$$\psi_{yt} + \psi_y\psi_{xy} - \psi_x\psi_{yy} = \frac{d}{dt}u_e(t) + \nu\psi_{yyy} \quad \ldots\ldots [1]$$

where subscripts denote partial differentiation, and ν is the kinematic-viscosity coefficient. The nonslip condition on the surface leads to the following two boundary conditions for

$$\psi_x = \psi_y = 0 \text{ at } x \geq 0 \text{ and } y = 0 \text{ for all } t \geq 0 \ldots [2a]$$

and the condition of continuity of u at the leading edge $x = 0$ gives

$$\psi_y = u_e(t) \text{ at } x = 0 \text{ and } y > 0 \text{ for all } t \geq 0 \ldots\ldots [2b]$$

To find a solution for small $\xi = x/\int^t u_e(t)dt$ we define, following the procedure in reference (1)

$$\xi_0 = \xi^{1/3}, \quad \eta = y(u_e'/\nu u_e)^{1/3}, \quad \zeta_0 = \eta/\alpha_0(t)\xi_0$$

and define

$$\psi = \theta(t)\Sigma F_r(\zeta_0, t)\xi_0^{r+1+\beta} \ldots\ldots\ldots\ldots [3]$$

Where s and β are undetermined constants; $\alpha_0(t)$ and $\theta(t)$ are undetermined functions of t.

The quantity η as well as $\alpha_0(t)$, defined in Equation [3], becomes imaginary if $u_e'(t) < 0$. This definition is adopted for convenience in preference to the ones defined in terms of the absolute value of $u_e'(t)$.[5] Summation will extend from 0 to ∞ if not otherwise specified. With this definition of ψ we have

$$\frac{u}{u_e} = [\theta(t)u_e'^{1/3}\alpha_0^{-1}\nu^{-1/3}u_e^{-1/3}]\Sigma\frac{\partial F_r}{\partial\zeta_0}\xi^{r+\beta}$$

Thus the initial condition [2b] requires that

$$\beta = 0, \quad \theta(t) = \alpha_0(t)\nu^{1/3}u_e^{2/3}u_e'^{-1/3} \ldots\ldots\ldots [4]$$

The constant s is found to be $s = 2$ in order to admit a nontrivial, consistent, downstream solution. For this initial approximation, $F_0(\zeta_0 t)$, the transient effects do not enter with convective effects alone balancing the viscous effects, but the differential equation for $F_0(\zeta_0, t)$ stands as

$$\frac{\partial^3 F_0}{\partial\zeta_0^3} + \left[\frac{\alpha_0^3}{2}u_e^2/u_e'\int_0^t u_e dt\right]F_0\frac{\partial^2 F_0}{\partial\zeta_0^2} = 0 \ldots\ldots [5]$$

even though the boundary conditions are all independent of t, given as

$$F_0(0, t) = \frac{\partial F_0}{\partial\zeta_0}(0, t) = 0, \quad \lim_{\zeta_0\to\infty}\frac{\partial F_0}{\partial\zeta_0}(\zeta_0, t) = 1 \ldots [6]$$

Equation [5] admits $F_0(\zeta_0, t)$ as a separable function of ζ_0 and t. But to satisfy Equation [6], the time-dependent part can only be a constant in agreement with physical intuitions. Thus the only permissible form of $\alpha_0(t)$ is

$$\alpha_0^3(t) = u_e'/u_e^2\int_0^t u_e dt \ldots\ldots\ldots\ldots [7]$$

except for a multiplying constant which is taken as unity for convenience so that $F_0(\zeta_0)$ may be identical with the Blasius solution (6, 7).

Summing up we have, for a solution around $\xi_0 = 0$

$$\xi_0 = \xi^{1/3} = \left[x\Big/\int_0^t u_e dt\right]^{1/3}, \quad \zeta_0 = \eta/\alpha_0\xi_0 = y[u_e/\nu x]^{1/3} \ldots [8]$$

and

$$\psi = (\nu u_e)^{1/3}\Sigma F_r(\zeta_0, t)\xi_0^{r+1}$$

The Boundary Conditions [2] then stand as

$$F_r(0, t) = \frac{\partial F_r}{\partial\zeta_0}(0, t) = 0 \quad \text{for} \quad t > 0, \quad \lim_{\zeta_0\to\infty}\frac{1}{\zeta_0^r}\frac{\partial F_r}{\partial\zeta_0} = \delta_{0,r} \ldots\ldots [9]$$

where $\delta_{0,r}$ is the Kronecker's delta which is zero for $r \neq 0$ and is unity for $r = 0$.

By substituting Series [8] into the Equation of Motion [1] and equating coefficients of like powers of ξ_0, a series of ordinary differential equations is obtained. Investigation of these equations

[5] Reference (2), appendix D.

shows that all the $F_r(\zeta_0, t)$ for r = odd integer vanish identically as a result of the behavior of the homogenous solution at large ζ_0.[6] The asymptotic behavior of $F_{2r}(\zeta_0, t)$ at large ζ_0 for $r \geq 1$ indicates that all the $F_{2r}(\zeta_0, t)$ are sums of separable functions of ζ_0 and t.[7] The separated functions of ζ_0 will be denoted by f with two subscripts; e.g., f_{62} means the second function in $F_6(\zeta_0, t)$. Single subscript means it involves no sum.

Investigation of the asymptotic behavior of the solutions of these equations shows that for the solution around the leading edge, the third boundary conditions of Equation [9] are identical with the condition that

$$\lim_{\zeta_0 \to \infty} \frac{\partial F_r}{\partial \zeta_0} = \delta_{0,r} \quad \ldots\ldots\ldots\ldots\ldots\ldots [10]$$

Thus, the solutions of these equations taken together will satisfy automatically the condition at the outer edge of the boundary layer $u/u_e = 1$. A number of these equations have been integrated numerically. The following values of the initial second derivatives are obtained as

$$\left.\begin{array}{ll} f_0''(0) = 0.33206 & f_2''(0) = 0.84851 \\ f_{41}''(0) = -0.46967 & f_{42}''(0) = 0.27088 \\ f_{81}''(0) = 0.36773 & f_{82}''(0) = -0.70141 \\ f_{83}''(0) = 0.37720 & f_{81}''(0) = -0.26949 \\ f_{83}''(0) = 0.84295 & f_{83}''(0) = 0.28896 \\ f_{84}''(0) = -1.68654 & f_{88}''(0) = 0.80413 \end{array}\right\} \ldots [11]$$

For detailed results, the reader is referred to reference (2).

We are particularly interested in the velocity profile u/u_e and the skin-friction coefficient $c_f = \tau_w \Big/ \frac{1}{2}\rho_e u_e^2$. These are obtained from the previous series solution, valid for sufficiently small values of ξ, as

$$\frac{u}{u_e} = f_0' + u_e' f_2' \frac{x}{u_e^2} + [u_e u_e'' f_{41}' + u_e'^2 f_{42}'] \frac{x^2}{u_e^4} + [u_e^2 u_e''' f_{81}' + u_e u_e' u_e'' f_{82}' + u_e'^2 f_{83}'] \frac{x^3}{u_e^6} + [u_e^2 u_e^{(iv)} f_{81}' + u_e^2 u_e' u_e''' f_{82}' + u_e^2 u_e''^2 f_{83}' + u_e u_e'' u_e'^2 f_{84}' + u_e'^4 f_{85}'] \frac{x^4}{u_e^8} + \ldots\ldots\ldots\ldots [12]$$

and

$$C_f R_{ex}^{1/2} = 2 \left\{ f_0''(0) + u_e' f_2''(0) \frac{x}{u_e^2} + [u_e u_e'' f_{41}''(0) + u_e'^2 f_{42}''(0)] \frac{x^2}{u_e^4} + [u_e^2 u_e''' f_{81}''(0) + u_e u_e' u_e'' f_{82}''(0) + u_e'^2 f_{83}''(0)] \frac{x^3}{u_e^6} + [u_e^2 u_e^{(iv)} f_{81}''(0) + u_e^2 u_e' u_e''' f_{82}''(0) + u_e^2 u_e''^2 f_{83}''(0) + u_e u_e''' u_e'^2 f_{84}''(0) + u_e'^4 f_{85}''(0)] \frac{x^4}{u_e^8} + \ldots\ldots\ldots\ldots [13]$$

It is seen that the final form of these Expressions [12] and [13] do not involve $\int_0^t u_e(t)dt$ explicitly, a result which could not have been anticipated from the case when $u_e \sim t^n$ as considered in reference (1).

It should be noted that the previous series solution could possibly "converge" over sufficiently small but finite values ξ only when all the coefficients have a finite upper bound. Since all the $f_r'(\zeta_0)$ have finite derivatives, it is clear that such series solution cannot be valid for any finite value of ξ (or x) when $u_e(t) = 0$ at some instant t, or when any of the timewise derivatives of $u_e(t)$ becomes infinite. An immediate inference from this is that this series solution will converge relatively more poorly in a decelerating period than in an accelerating period, and is totally useless even for engineering purposes, during the period of time when $u_e(t)$ is decelerated to a very small value.

The fact that the Expressions [12] and [13] are independent of $\int_0^t u_e dt$ indicates that they are the proper limiting forms for large t or $\int_0^t u_e dt$ provided that these series converge. While the present solution is valid only for unsteady motions started from rest continuously (as will be shown later), it seems that this limiting form for large t may be a valid approximation even for unsteady motions not started from rest, because the starting conditions would leave little effect on the flow field when the unsteady motion has proceeded for sufficiently long periods of time. For example, the large time behavior of the flow field induced by a plate started impulsively from rest should be substantially the same as that induced by the motion of a plate described by $u_e = U[1 - \exp(-kt)]$ where k is a positive constant. As an additional example, the periodic flow over an oscillating plate in a uniform stream may be considered as the large time behavior of the flow field induced by a plate motion $u_e = U[1 - \exp(-kt)][1 + a \sin \omega t]$ where a is some positive constant less than unity. Computations of the two cases mentioned here have been carried out. The solution is valid (as will be seen later) for all time within the range of values of x as to have Equations [12] and [13] converge. The velocity profiles are plotted in Figs. 1 and 3 and the skin friction in Figs. 2 and 4. At a given time, the velocity profile is seen to approach the Rayleigh-type profile as ξ increases; the skin friction also approaches the Rayleigh value as ξ increases, but eventually, for sufficiently large values of ξ, the Series [12] and [13] (with the five terms determined so far) will diverge and are no longer applicable. In order to determine the flow field in the downstream region for practical engineering purposes and to investigate its asymptotic behavior in the far downstream end for an analytic justification of the present solution, it is necessary to continue the previous series solutions into the downstream region.

Solutions for $\xi > \xi_i$

Let us now consider the continuation of the solution from an initial station $\xi = \xi_i$ where the local velocity profile $u/u_e(\eta, t)$ is known, into the region $\xi > \xi_i$. The method of continuation is a natural extension of the method used in reference (1) and is divided into three parts. A series solution is determined from the known u/u_e for small η and the nonslip condition on the surface $\eta = 0$ with the technique used near the leading edge. This series solution is valid only near the wall since the boundary condition at the outer edge of the boundary layer is no longer satisfied automatically as in the preceding section. For large values of η near the outer edge of the boundary layer, a perturbation solution over the initial profile as suggested by the solution for small η is constructed. This perturbation solution is found to have a logarithmic singularity at a point η_i where $u/u_e(\xi_i, \eta_i, t) = \xi_i$ and if extended to the region of small η is found to agree with the local series solution about ξ_i if the rearrangement of terms in the series is permitted. The solution in the vicinity of $\eta \sim \eta_i$ is then constructed independently to join the solution for large η and that for small η.

The series solution valid near the wall for $\xi > \xi_i$ is determined

[6] Reference (2), appendix A.
[7] Reference (2), appendix B.

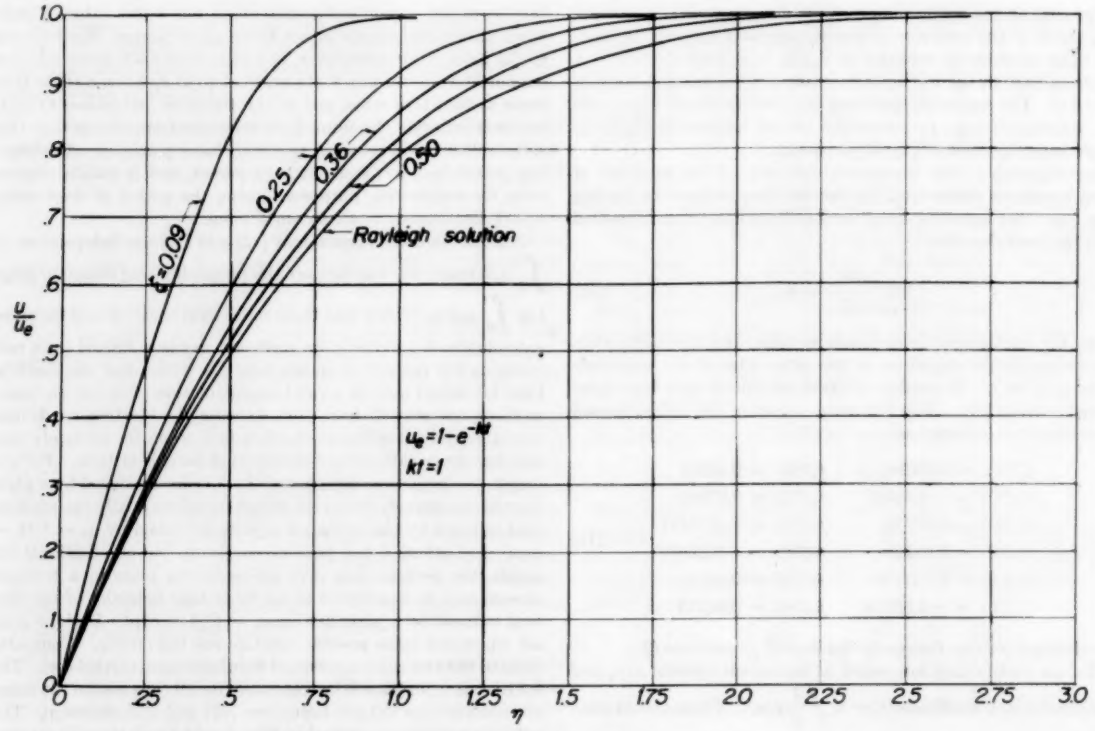


FIG. 1

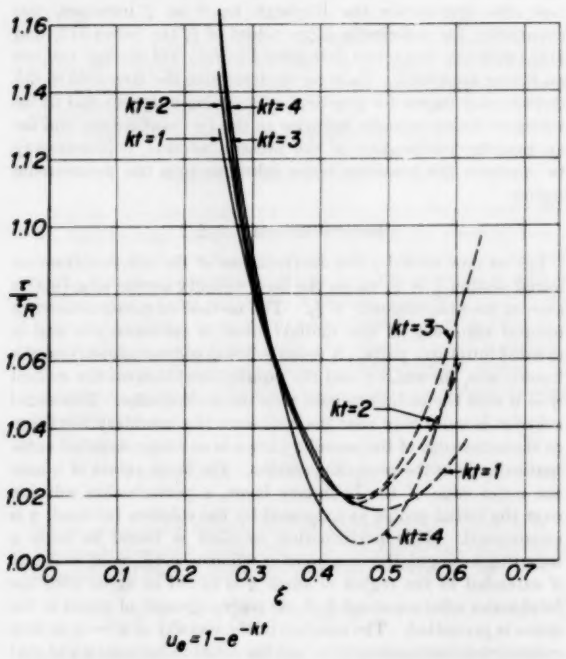


FIG. 2

with the following series expansion as a result of an analysis similar to the one given in the preceding section.

$$\psi = \left[2\nu u_e \xi_i^{-1} \int_0^t u_e dt\right]^{1/2} \Sigma\, C_{r+1}(\xi_i, t) F_r(\zeta_1, t) \alpha_1^{r+1} \xi_1^{r+2}$$

with

$$\xi_1 = (\xi - \xi_i)^{1/2}, \quad \eta = y(u_e'/\nu u_e)^{1/2},$$

$$\alpha_1^2(t) = 2u_e' \int_0^t u_e dt / \xi_i u_e^2 \ldots\ldots [14]$$

and

$$\zeta_1 = \eta/\alpha_1\xi_1 = y\left[u_e/2\nu\left(x/\xi_i - \int_0^t u_e(t)dt\right)\right]^{1/2}$$

where the initial velocity profile $u/u_e(\eta, t)$ is given as

$$\frac{u}{u_e} = \Sigma\, C_{r+1}(\xi_i, t)\eta^{r+1} \ldots\ldots [15]$$

With the series expansion [14], the convective terms enter as higher order small terms so that the balance is primarily between the transient effects and the viscous effects, a situation to be expected near the wall and quite different from the expansion [8]. It is natural to expect that Series [14] will take over the Series [8] at smaller values of ξ_i, if the transient effects are relatively more significant as compared with convective effects.

The coefficients $C_r(\xi_i, t)$ are not all independent because u/u_e must satisfy the Equation of Motion [1] with $\psi_y = u$. The following relations must hold with $C_0 = 0$ and $C_2 = -1/2$

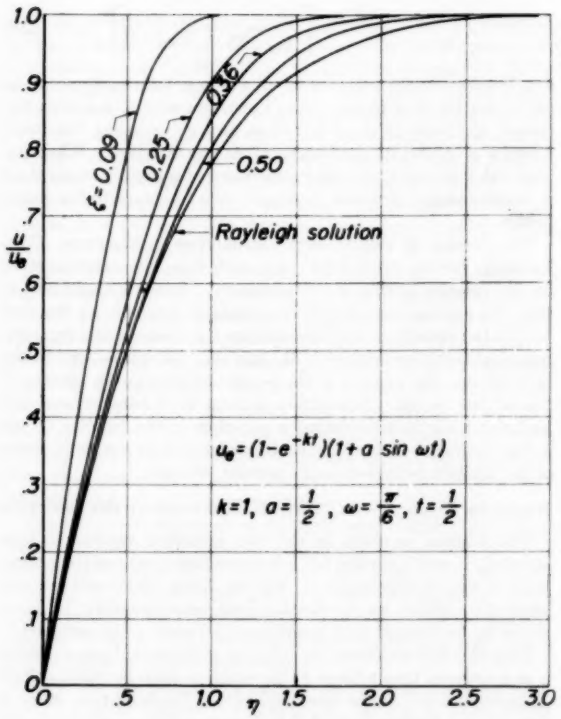


FIG. 3

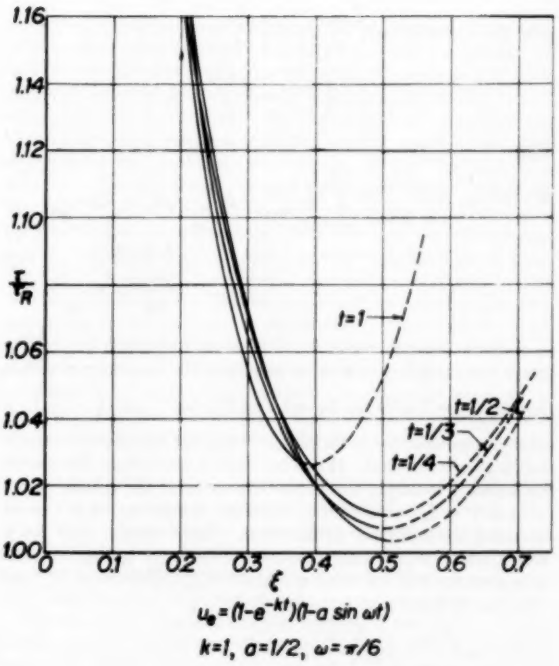


FIG. 4

$$(l+2)(l+1)C_{l+2} = -\delta_{0,l} + C_l + \frac{u_e}{u_e'}\frac{\partial C_l}{\partial t} + \frac{1}{2}\left(\frac{u_e u_e''}{u_e'^2} - 1\right) l C_l - \frac{u_e^2 \xi_i}{u_e' \int_0^t u_e dt}\frac{\partial C_l}{\partial \xi_i} + \frac{u_e^2}{u_e' \int_0^t u_e dt}\sum_{k=1}^{l-1}\frac{l-2k+1}{l-k+1} C_k \frac{\partial C_{l-k}}{\partial \xi_i} \quad \ldots\ldots [16]$$

With the relations [16] the functions $F_r(\zeta_1, t)$ satisfying the boundary conditions

$$F_r(0,t) = \frac{\partial F_r}{\partial \zeta_1}(0,t) = 0$$

$$\lim_{\zeta_1 \to \infty} \frac{1}{\zeta_1^{r+1}}\frac{\partial F_r}{\partial \zeta_1} = 1 \quad \ldots\ldots [17]$$

are $F_n(\zeta_1, t) = \dfrac{\zeta_1^{n+2}}{n+2} + \dfrac{1}{C_{n+1}}\sum_{r=1}^{[n/2]}\dfrac{1}{\alpha_1^{2r} r!}\dfrac{\partial^r C_{n-2r+1}}{\partial \xi_i^r}\dfrac{\zeta^{n-2r+2}}{n-2r+2} \quad \ldots\ldots [18]$

where $[n/2]$ denotes the integral part of $n/2$.

By substituting Equation [18] into [14] and differentiating with respect to y to obtain $u = \psi_y$, we find that the resulting series for u/u_e involves only even powers of ξ_1; in other words, it is a regular function at and about ξ_i. This suggests that the profile in the vicinity downstream of an initial station ξ_i can be determined by regular perturbation procedures[3] provided, of course, the rearrangement of terms in the previous series solutions can be carried out. Accordingly we look into the possibility of constructing the downstream solution by perturbing the known solution at ξ_i. Thus let

$$\psi = u_e^{3/2} u_e'^{-1/2} \nu^{1/2} f(\xi, \eta, t) = \nu^{1/2} u_e^{3/2} u_e'^{-1/2} \Sigma g_r(\eta, t) \alpha_1^{2r} (\xi - \xi_i)^r \quad \ldots\ldots [19]$$

This form of expansion is required by the joining of this solution with that for small η. For a fixed (η, t), and letting $\xi \to \xi_i$, we have

$$\partial g_0/\partial \eta = (u/u_e)_{\xi=\xi_i}$$

which is a known function. This ensures that the velocity profiles match at $\xi = \xi_i$. The boundary condition at the outer edge of the boundary layer that $u/u_e = 1$ is satisfied by $\partial g_0/\partial \eta = 1$ as $\eta \to \infty$. Therefore all the g_k's with $k \geq 1$ will be required to satisfy

$$\lim_{\eta \to \infty} \frac{\partial g_k}{\partial \eta} = 0 \quad \text{with } k \geq 1 \quad \ldots\ldots [20]$$

By substituting Equation [19] into Equation [1] and equating the like powers of $\xi - \xi_i$ we obtain a series of first-order inhomogeneous differential equations for g_k which can be required to satisfy one boundary condition at some value of η at our disposal. Consider the first-order perturbation $g_1(\eta, t)$ satisfying

$$\left(\frac{\partial g_0}{\partial \eta} - \xi_i\right)\frac{\partial g_1}{\partial \eta} - \frac{\partial^2 g_0}{\partial \eta^2} g_1 = \frac{\xi_i}{2} I_1(\xi_i, \eta, t)$$

where $I_1(\xi_i, \eta, t) = 1 - \dfrac{\partial g_0}{\partial \eta} + \dfrac{\partial^2 g_0}{\partial \eta^2} - \dfrac{u_e}{u_e'}\dfrac{\partial^2 g_0}{\partial \eta \partial t} - \dfrac{1}{2}\left(\dfrac{u_e u_e''}{u_e'^2} - 1\right)\eta \dfrac{\partial^2 g_0}{\partial \eta^2} \quad \ldots\ldots [21]$

The solution is

$$g_1(\eta, t) - g_1(\infty, t) = \left(\frac{\partial g_0}{\partial \eta} - \xi_i\right)\frac{\xi_i}{2} \int_\infty^\eta \left(\frac{\partial g_0}{\partial \eta} - \xi_i\right)^{-2} I_1(\xi_i, \eta, t)d\eta \quad \ldots\ldots [22]$$

with

$$\frac{\partial g_1(\eta, t)}{\partial \eta} = \frac{\xi_i}{2}\left\{\frac{\partial^2 g_0}{\partial \eta^2}\int_\infty^\eta \left(\frac{\partial g_0}{\partial \eta} - \xi_i\right)^{-2} I_1(\xi_i, \eta, t)d\eta + \frac{I_1(\xi_i, \eta, t)}{\dfrac{\partial g_0}{\partial \eta} - \xi_i}\right. \quad \ldots\ldots [23]$$

a similar expression can be obtained easily for g_k with $k > 1$. The lower limit is taken as $\eta = \infty$ to satisfy the boundary condition $\lim \partial g_1/\partial \eta = 0$ with $\partial g_0/\partial \eta \to 1$ and $\partial^2 g_0/\partial \eta^2 \sim \exp\left(-\frac{\eta^2}{4}\right)$. A similar argument holds for all g_k. Thus the boundary condition for large η is satisfied. It can be verified easily that this g-series formulation is simply the Taylor expansion of the velocity about u/u_e at $\xi = \xi_i$ provided that the order of summation can be reversed without further justification. This indicates that the g-series, which is primarily constructed for large η, may be valid also near the wall for small η. Accordingly, Equations [22] and [23] can be written as

$$g_1(\eta, t) = \left(\frac{\partial g_0}{\partial \eta} - \xi_i\right)\frac{\xi_i}{2} \int_0^\eta \left(\frac{\partial g_0}{\partial \eta} - \xi_i\right)^{-2} I_1(\xi_i, \eta, t)d\eta \quad \ldots\ldots [24]$$

$$\frac{\partial g_1}{\partial \eta}(\eta, t) = \frac{\xi_i}{2}\left\{\frac{\partial^2 g_0}{\partial \eta^2}\int_0^\eta \left(\frac{\partial g_0}{\partial \eta} - \xi_i\right)^{-2} I_1(\xi_i, \eta, t)d\eta + \frac{I_1(\xi_i, \eta, t)}{\dfrac{\partial g_0}{\partial \eta} - \xi_i}\right. \quad \ldots\ldots [25]$$

It is verified that Equation [25] gives the same skin friction as does the previous series solution near the wall using the ξ_1 or f-series. Thus the g-series with g_1 and $\partial g_1/\partial \eta$ determined by Expressions [22] to [25] and similar expressions for g_k and $\partial g_k/\partial \eta$ with $k > 1$ would complete the continuation of the solution into downstream regions if the integrals were proper for all values of η. This would be the case when ξ_i is sufficiently large. But, for practical purposes, the continuation would have to start from $\xi_i < 1$. Since $\partial g_0/\partial \eta$ varies from 0 to 1 when η increases from 0 to ∞, the integrals of Equations [23] and [25] for the determination of the velocity u are logarithmically singular when $\eta = \eta_i$ with $\partial g_0/\partial \eta\,(\eta_i, t) = \xi_i$. Thus Equations [22] and [23] can only be valid for large $\eta > \eta_i$, and [24] and [25] for small $\eta < \eta_i$. Numerical evidence is that this logarithmic singularity is very "weak." The profile as evaluated from Equation [25] is compared with the series solution around the leading edge when it is still valid and is found to be not distinguishable for practically all values of $\eta < \eta_i$. However, the distortion of u/u_e due to the logarithmic singularity is quite significant when u/u_e is evaluated from Equation [23] for say $2\eta_i > \eta > \eta_i$. Therefore, if one should be interested in the detailed profile of u/u_e, a solution valid in the vicinity of $\eta \geq \eta_i$ must be constructed.

It can be deduced from Equation of Motion [1] that

$$\left[\frac{\partial u}{\partial x}\Big/\frac{\partial u}{\partial y}\right]_{y_0} = \int_{y_0}^{y} \frac{1}{u^2}\left[\nu u_{yy} + u_e'(t) - u_t\right]dy + \left[\frac{1}{uu_y}\nu u_{yy} + u_e' - u_t\right]_{y_0}^{y} \quad \ldots\ldots [26]$$

if y_0 is taken in the region $\eta < \eta_i$, where a sufficiently accurate determination of u in the region downstream of ξ_i has been obtained, the local value of $\partial u/\partial x$ at any y (including the point where $\eta = \eta_i$ and its neighborhood) can be obtained. When the local value of u at η_i or rather some value of η slightly larger than η_i is determined, it would be simple then to complete the entire profile.

The process of step-by-step continuation is laborious. Fortunately, for the present purpose, such complete determination of the velocity profile is not necessary. What is important is that the g-series formulation is capable of determining the flow field in the vicinity of the plate despite the presence of a logarithmic singularity somewhere in the stream. As will be seen in the next section, the validity of the g-series at least in the vicinity of the surface in the continuation process is of fundamental importance both in establishing a criterion of the validity of the series expansion around the leading edge and in the verification of the limiting behaviors of the present solution.

Range of Validity and Limiting Behaviors of the Solution

The solution as given in the two preceding sections is constructed, based upon the boundary conditions (i) and (ii) as outlined in the "Introduction." The condition (iii) is either automatically satisfied, or the solution is adjusted to satisfy it. It remains to be verified that conditions (iv) and (v) be satisfied.

Consider first condition (iv) that for sufficiently large values of x at any given time t (large ξ) the solution must be substantially independent of ξ and is described by the Rayleigh-type solution $G(\eta, t)$. By putting $\partial/\partial\xi = 0$ in Equation [1] we obtain the following partial differential equation for $G(\eta, t)$

$$1 - \frac{\partial G}{\partial \eta} + \frac{\partial^3 G}{\partial \eta^3} + \frac{\eta}{2}\frac{\partial^2 G}{\partial \eta^2} = \frac{u_e}{u_e'}\frac{\partial^2 G}{\partial \eta \partial t} + \frac{1}{2}\frac{u_e u_e''}{u_e'^2}\eta\frac{\partial^2 G}{\partial \eta^2} \quad \ldots\ldots [27]$$

Subject to the boundary conditions

$$G(0, t) = \frac{\partial G}{\partial \eta}(0, t) = 0, \quad \lim_{\eta \to \infty}\frac{\partial G}{\partial \eta} = 1 \quad \text{for all } t \geq 0$$

The solution for velocity is (8)

$$u_R(y, t) = u_e(t)\frac{\partial G}{\partial \eta}(\eta, t) = \int_0^t u_e'(t - \sigma)\,\mathrm{erf}\left(\frac{y}{2\sqrt{(\nu\sigma)}}\right)d\sigma + u_e(0)\,\mathrm{erf}\left(\frac{y}{2\sqrt{(\nu\sigma)}}\right) \quad [28]$$

The shear stress at the wall is

$$\tau_R(0, t) = \mu\left[\frac{\partial}{\partial y}u_R(y, t)\right]_{y=0} = \left[\int_0^t u_e'(t - \sigma)\sigma^{-1/2}d\sigma + \frac{u_e(0)}{t^{1/2}}\right]\left(\frac{\rho\mu}{\pi}\right)^{1/2} \quad \ldots\ldots [29]$$

Our problem is to see if the present solution will approach $G(\eta, t)$ as ξ increases.

Let us first note from Equation [19] that

$$\alpha_1^2\frac{\partial g_1}{\partial \eta} = \left[\frac{\partial}{\partial \xi}(f_\eta)\right]_{\xi=\xi_i} = \left[\frac{\partial}{\partial \xi}\left(\frac{u}{u_e}\right)\right]_{\xi=\xi_i} \quad \ldots [30]$$

The sign of $\partial g_1/\partial\eta$ as given by Equation [25] will therefore determine whether the local velocity u/u_s increases or decreases with ξ increasing from ξ_i. With $I_1(\xi_i, \eta, t)$ defined in Equation [21] it is immediately seen that if $g_0(\xi_i, \eta, t)$ is identical with $G(\eta, t)$, I_1 will vanish identically for all η and t; in other words, $\alpha_1{}^2\partial g_1/\partial\eta = (\partial u/\partial\xi)_{\xi=\xi_i}$ vanish for all η and t, and the velocity profile will not undergo any change once the Rayleigh-type solution is reached. Moreover, the integral in Equation [25] is from $\eta = 0$ to η. Thus the skin friction will remain unchanged while proceeding downstream once the local skin friction reaches the Rayleigh value given by Equation [29].

For sufficiently small values of η, Equations [25] and [30] indicate that $\partial u/\partial\xi$ at small η and ξ_i is proportional to $-\alpha_1{}^2 I_1$. Therefore $\partial u/\partial\xi$ can vanish only when I_1 is zero; that is, only when the skin friction is given by the Rayleigh. Now in the region with sufficiently small values of ξ where the series expansion about the leading edge is valid, $\alpha_1{}^2 I_1$ is positive and $\partial u/\partial\xi$ negative with their values decreasing with increasing ξ. Therefore, the skin friction of the present solution, when properly continued, will approach monotonically the Rayleigh value from the positive side with increasing ξ and the solution will be substantially Rayleigh at sufficiently large values of ξ.

The continuation process certainly should be valid at any small values of ξ away from the leading-edge singularity. The previous result that $\partial u/\partial\xi$ is proportional to $-\alpha_1{}^2 I_1$, which vanishes only when the Rayleigh solution is reached, leads to a criterion for the determination of the largest possible value of ξ_i within which range the f-series solution may serve as a valid approximation with a finite number of terms in the series.

When the calculated skin friction reaches the Rayleigh value τ_R with a finite slope $\partial\tau/\partial\xi$ or reaches a minimum where $\partial\tau/\partial\xi$ vanishes at a value of τ_{min} larger than τ_R, the f-series representation is already invalid. For a prescribed accuracy, one can determine the station ξ_i where the f-series representation is still satisfactory by redetermining the profile or the skin friction at ξ_i using the g-series starting from the profile at a slightly upstream station ξ_{i-1}. For the cases with $u_s \sim t^n$ (reference 1) for different positive values of n, the five-term representation is satisfactory up to $\xi_i = 0.4$ where the skin friction is only 2 per cent above the Rayleigh value τ_R. For the cases with $u_s \sim 1 - \exp(-kt)$ the calculated results with five terms in Series [13] is given in Fig. 2. The value of ξ_i is again $\sim$0.4 with τ at $\xi = 0.4$ about 3 per cent above τ_R. For the cases with $u_s \sim [1 - \exp(-kt)]\cdot[1 - a \sin \omega t]$ the situation is rather similar to the two previous examples when ωt takes values in the accelerating range. When ωt takes values in the decelerating range, the accuracy of Series [13] with five terms is considerably poorer, so that ξ_i would have to be taken less than 0.4 with the local skin friction deviating more than a couple of per cent from τ_R. Under these circumstances, it probably would be necessary to carry out one or more steps of continuation process with g-series in order to bring the local skin friction τ to within prescribed error from τ_R for different engineering purposes. The skin friction along the surface downstream of that station may as well be taken as τ_R or with an arbitrarily faired-in value toward τ_R.

The calculated examples indicate that the Rayleigh-type solution is in fact a very good engineering approximation, at least in so far as skin friction is concerned for $\xi_i \geq 0.5$. This rough overall picture may be expected to be quite good for accelerating flow. When the flow is decelerating this rough estimate may not be so good and would better be investigated more carefully as just outlined for each specific motion.[9]

[9] For decelerating flow τ_R may reach very small magnitude which results in large fractional deviation of τ/τ_R from unity. The absolute error, however, may not be very significant.

Now we can investigate the behavior of the solution in the limit of very small time. Consider the limiting forms, Equations [12] and [13], of the series solution around the leading edge. The series can possibly converge only when x vanishes as $u_s{}^{1/2}$, if $u_s{}'$, and so on, remain finite in the limit. Therefore, they will apply only to a vanishingly small region around the leading edge, immediately followed by the continued solution; i.e., by the Rayleigh-type solution. Thus, at very small t the flow field is practically Rayleigh over the entire plate, as expected physically when the plate started to move from rest. It is now clear that the present solution certainly does not apply to unsteady motions of plate not started from rest, because at $t = 0$ we would expect in this case the flow field to be Blasius all the way around the plate corresponding to the steady-state velocity $u_s(0)$ in the limit of $t = 0$.

Of course, it should be borne in mind that the boundary-layer approximation cannot be relied upon for even an approximate description of the flow field too close to the leading edge where R_{xs} is not $\gg 1$. The result, Equation [12], does not imply that, for sufficiently small values of x, the local flow condition is actually quasisteady.

As was pointed out in a previous section, the result, Equation [12], may be a good approximation in the limit of large times, as physically appears necessary even if the unsteady motion does not start from rest. But Equation [12] cannot be expected to determine the deviation from the large time behavior at a given finite time when the unsteady motion did not start from rest.

To sum up, the present solution consists of two parts, a series solution around the leading edge valid for small x so long as the f-series, Equations [12] and [13], converge. Beyond the range of validity of these series, the stepwise continuation procedure with the g-series, Equations [19] and [25], will determine the downstream flow field, if ever required. The range of validity of the f-series, Equations [12] and [13], can be established by comparing the predicted profiles at the same station with the f-series and the g-series. From this station downstream, the skin friction must approach the Rayleigh value τ_R given by Equation [29] monotonically. If, within the range of validity of the Series [12] and [13] τ is sufficiently close to τ_R one can take from this station downstream the skin friction to be given by the τ_R. For most of accelerating unsteady flow, the error involved by taking $\tau = \tau_R$ for $\xi \geq 0.5$ is not more than 2 or 3 per cent.

Comparison With Results of Lighthill and Moore

Experimental data are as yet not available to check the results in this paper. Therefore, a comparison with analytic results of other authors is attempted.

Lighthill (4) has investigated unsteady flow over a general two-dimensional body when there is a small oscillation in the magnitude of the external stream. Lighthill writes $u_s = U(x)(1 + ae^{i\omega t})$ where $|a| \ll 1$. By an ingenious use of the momentum-integral technique the skin friction and the velocity profile were found. Under the further assumption of $\omega l/U \ll 1$ where l is some characteristic length, Lighthill gives the following general results

$$\tau_w = \tau_0 + \text{Real Part of } \left[ae^{i\omega t}\left(\frac{3}{2}\tau_0 + i\omega\frac{\rho}{2}U\delta_0{}^*\right)\right] \ldots [31]$$

Where τ_0 and $\delta_0{}^*$ are the skin friction and the displacement thickness of the quasi-steady flow, respectively. In the particular case of the flat plate, this reduces to

$$\tau_w = \mu U\sqrt{\left(\frac{U}{\nu x}\right)}\left\{0.33206 + a\left[\frac{3}{2}\times 0.33206\cos\omega t - 0.8604\frac{\omega x}{U}\sin\omega t\right]\right\} \ldots\ldots [32]$$

Now from the f-series solution around the leading edge τ_w is given by Equation [13] for oscillating flows over a flat plate but not necessarily of small amplitudes. If we substitute $u_e = U(1 + ae^{i\omega t})$ into Equation [13] and neglect terms of the order of a^2 with Lighthill's assumption, we have

$$\tau_w = \mu U \sqrt{\left(\frac{U}{\nu x}\right)} \left\{0.33206 + a\left[\frac{3}{2} \times 0.33206 \cos \omega t - 0.84851 \frac{\omega x}{U} \sin \omega t - 0.46967 \left(\frac{\omega x}{U}\right)^2 \cos \omega t + 0\left(\frac{\omega x}{U}\right)^3\right] + 0(a^2)\right\} \quad [33]$$

It is surprising to find that Equations [32] and [33] agree remarkably not only in form but also in the numerical constants. This illustrates very well the potentiality of the intelligent use of the momentum-integral technique.

Even more surprising is the deep physical insight of Lighthill's proposal of joining his solution at low frequency ω where $\omega l/U \ll 1$ and the one at high frequency where $\omega l/U \gg 1$. He assumes that the boundary-layer solutions can be approximated by the "shear-wave" solution where the inertia terms are completely neglected at high frequencies. This is precisely what we call the Rayleigh-type solution in which the skin-friction fluctuation has a phase advance of $\pi/4$ over the free-stream fluctuation for all values of ω. Lighthill assumed that for all frequencies ω greater than a critical frequency ω_0, the solution should be given by the shear-wave solution (or the Rayleigh-type solution). The critical frequency ω_0 is determined from the low-frequency solution Equation [32] at which the skin-friction fluctuation leads the free-stream fluctuation by $\pi/4$. From the point of view taken in this paper, Equation [13] is valid only for those values of $\omega x/U$ less than a certain critical magnitude to be determined as explained at the end of the preceding section. For $\omega x/U$ larger than this critical value, the solution can be continued as described, or can be faired in arbitrarily or taken as the Rayleigh-type solution depending upon circumstances and the accuracy required. Lighthill's assumption concerning the determination of ω_0 is, in a sense, corresponding to the process of extrapolating the skin-friction curve as determined from Equation [33] to the Rayleigh value and of taking the Rayleigh value for all the downstream stations.

Lighthill's analysis is more general than the present one in the aspect that it considers various geometries of thin bodies but is more restricted in the aspects of being interested only in small perturbations and in flow fields over bodies of finite lengths. It is the last two restrictions that enable him to define a critical frequency ω_0 for matching, which probably can be more accurately determined with the Series [33]. In the case of nonlinear oscillations, the present analysis indicates that the parameter for determining the matching station is $\omega x/U$ rather than the frequency ω only. The critical frequency ω_0 in this case will depend very much on the relative amplitude of the oscillation, the distance x downstream of the leading edge as well as the instant t under consideration. The extension of the present analysis to arbitrary thin bodies as considered by Lighthill does not seem to present any unexpected difficulty.

Let us now consider the analysis of Moore (5). Based upon dimensional arguments, he conceived an expansion of the stream function in terms of an infinite number of parameters defined by xu_e'/u_e^2, x^2u_e''/u_e^3 . . . ., etc., which he arranged in a sequence of "descending magnitudes." The series finally obtained is, in fact, identical with the f-series solution determined in the present paper for solutions around the leading edge except for a quite thorough rearrangement of the terms. The coefficients of the three terms which Moore determined agree with the corresponding ones in this paper. He was not primarily interested in the question of validity of the solution, the investigation of the limiting behaviors, and the requirement of matching. He is more interested in applying his result to a variety of physical problems, some of which the present analysis does not seem to justify even though no positive statement can be made. Problems involving unsteady motions not started from rest and those involving flow reversal require more careful investigation.

Summary and Conclusions

The problem of a boundary layer on a flat plate moving with an arbitrary velocity into an incompressible fluid at rest is investigated. A solution is constructed when the plate velocity $u_e(t)$ is differentiable, without flow reversal, and started from rest. This solution satisfies the following conditions:

(i) Nonslip condition of the plate surface.

(ii) The condition of uniform flow at the leading edge.

(iii) The condition of uniform stream far away from the surface.

(iv) At very large distance downstream of the leading edge, the solution is substantially independent of x and is practically the "shear-wave" or the "Rayleigh-type" solution.

(v) At very small time (the beginning of the unsteady motion), the entire flow field is nearly stagnant and the flow field in the vicinity of the plate is given by the Rayleigh-type or the shear-wave solution.

The solution consists of two parts:

(i) A series solution about the leading edge with the velocity and the skin friction given as Equations [12] and [13], respectively, valid for sufficiently small values of ξ.

(ii) When the Series [12] and [13] tend to diverge or are not sufficiently accurate downstream of $\xi = \xi_i$, the solution can be continued with Series [19] using [23], [25], and [26] for the three subregions in question.

Investigation of the continuation reveals that the proper continuation of the solution must approach the Rayleigh-type solution monotonically. Thus ξ_i must be upstream of the station where Series [13] either gives a skin friction equal to the Rayleigh value τ_R or reaches a minimum larger than τ_R. The range of validity of Series [12] and [13] may be more precisely determined for a prescribed accuracy by determining the skin friction and/or the velocity profile at ξ_i by using the continuation scheme starting from a slightly upstream station and comparing with those obtained directly from Series [12] or [13]. This depends very much on the particular type of unsteady motion $u_e(t)$ and the particular instant t. The following engineering rule may be adopted:

Series [13] may be used to determine the skin friction to or as far as it carries before the calculated skin friction either reaches the Rayleigh value or reaches a minimum.

1 If it reaches τ_R one may take the skin friction over the downstream stations as τ_R.

2 If it reaches a minimum value τ_{min} which is substantially the same as τ_R, one may fair the skin-friction curve, extrapolated to τ_R and take τ_R from this station downstream.

3 If it reaches a minimum value which is still significantly larger than τ_R and if the inaccuracy of fairing curves cannot be tolerated, then the process of continuation downstream as shown in the section, Solutions for $\xi > \xi_i$, becomes necessary until the required accuracy is obtained.

Sample calculations show that, for accelerating unsteady motion, Series [13] will lead to a minimum τ_{min} in the vicinity of $\xi \sim$

0.4 with τ_{min} only a couple of per cent above τ_R and will be a good description of the flow field for most engineering purposes. For decelerating periods of unsteady motions no such simple rule can be given and each individual case must be investigated carefully.

Comparison with Lighthill's solution for oscillating bodies in free streams indicates that the critical frequency ω_0 as defined by Lighthill may be generalized to $(\omega x/U)_0$ as the matching point of the upstream solution with the Rayleigh-type or the shear-wave solution.

BIBLIOGRAPHY

1 "Some Aspects of Unsteady Laminar Boundary Layer Flows," by S. I. Cheng, Princeton University, Aeronautical Engineering Department Report No. 311, July, 1955; to appear in January, 1957, issue of the *Quarterly of Applied Mathematics*.

2 "The Unsteady Laminar Boundary Layer on a Flat Plate," by S. I. Cheng and D. Elliott, Princeton University, Aeronautical Engineering Department Report No. 318, 1955.

3 "On the Impulsive Motion of a Flat Plate in a Viscous Fluid," by K. Stewartson, *The Quarterly Journal of Mechanics and Applied Mathematics*, vol. 4, part 2, June, 1951, pp. 182–198.

4 "The Response of Laminar Skin Friction and Heat Transfer to Fluctuations in the Stream Velocity," by M. J. Lighthill, Proceedings of the Royal Society of London, England, series A, vol. 224, no. 1156, June, 1954, pp. 1–23.

5 "Unsteady Laminar Boundary Layer Flow," by F. K. Moore, NACA TN 2471, 1951.

6 "Modern Development in Fluid Dynamics," by S. Goldstein, Clarendon Press, Oxford, England, vol. 1, 1938, pp. 181–190.

7 "Grenzschicht Theorie," by H. Schlicting, Verlag G. Braun, Karlsruhe, Germany, 1951.

8 "Unsteady Laminar Boundary Layer Over a Flat Plate in the Far Downstream Region," by S. I. Cheng and I. D. Chang, Princeton University, Aeronautical Engineering Department Report No. 339, 1956.

Ebullition From Solid Surfaces in the Absence of a Pre-Existing Gaseous Phase[1]

By S. G. BANKOFF,[2] TERRE HAUTE, IND.

The nucleation theory of Volmer and Fisher is extended to the superheating of liquids in contact with various solid boundaries. It is shown that nucleation in the bulk phase, at flat surfaces, at sharp projections, or in wetted cavities can be dismissed from consideration as possible explanations for experimentally observed superheats. Unwetted cavities are preferred nucleation points; but the difficulties of filling these completely with liquid are so great that it is probable that nucleation almost always occurs at a pre-existing gaseous phase in static stressing of the liquid. Nucleation in dynamic-stress phenomena, such as ebullition, cavitation, or effervescence, always occurs at the boundaries of gas or vapor entrapped in surface cavities.

Nomenclature

The following nomenclature is used in the paper:

A = area
C_{sf} = Rohsenow proportionality constant (1)[3]
$f(\theta)$ = function defined by Equation [4]
Δf_s^* = free energy of activation for evaporation of a single molecule (Equation [8])
h = Planck constant
k = Boltzmann constant
n = number of bubbles per mole of liquid
N = Avogadro number
P = pressure
ΔP = superheat-vapor-pressure difference
r = radius
R = gas constant
t = time
T = temperature, absolute
v = molal volume
V = volume
W = isothermal reversible work of formation
α, β = angles defined by Fig. 2 or Fig. 4
θ = contact angle measured through liquid
λ = molal latent heat of evaporation
μ = chemical potential
ρ = density
σ = interfacial tension
ϕ, ϕ' = function defined by Equation [15] or [Equation [17]

Superscripts

$'$ = vapor embryo
$*$ = vapor embryo of critical radius (i.e., in equilibrium with the surrounding liquid)

Subscripts

L = liquid
G = gas or vapor
S = solid

Introduction

When a superheated liquid nucleates to form vapor bubbles under essentially isopiestic conditions, this phenomenon is known as ebullition. If nucleation to form vapor bubbles occurs when the local pressure is reduced essentially isothermally below the normal vapor pressure of the liquid, this phenomenon is known as cavitation. In both cases nucleation is greatly facilitated if solid surfaces are present, to the extent that the nature and condition of these surfaces, in conjunction with the liquid-phase variables, usually control the nucleation process. Because of the complex geometry and surface chemistry of most actual surfaces, it has been quite difficult to predict the magnitude of the driving forces, either temperature or pressure, which are necessary for nucleation to occur in a specific solid-liquid combination. This complexity constitutes a major barrier to further advance in the fields of both ebullition and cavitation. For example, perhaps the most widely accepted correlation of nucleate boiling data, due to Rohsenow (1), predicts the slope of a log-log plot of heat flux versus temperature excess quite satisfactorily, but contains a proportionality constant, C_{sf}, which is an empirical function of the fluid-surface combination and which may vary over a six to ten-fold range. If it were possible to predict the temperature excess required to initiate nucleate boiling accurately, this would fix C_{sf}. Similar uncertainties exist in the field of cavitation.

At the same time there exists a wide diversity of opinion concerning the conditions necessary for bubble nucleation from solid surfaces. A number of investigators (2, 3, 4) are of the opinion that nucleation occurs only at the surfaces of minute quantities of vapor or gas entrapped by the solid. Jakob (5) considers that bubbles may form either at surface roughnesses or at entrapped gas surfaces. Larson (6) considers that the work of solid-liquid fracture, given by the Dupré expression in terms of the interfacial free energies and the contact angle

$$W_{LS} = \sigma_{LG} + (\sigma_{SG} - \sigma_{LS}) = \sigma_{LG}(1 + \cos\theta) \ldots\ldots [1]$$

determines the amount of superheat which a particular ebullator will support. If $W_{LS} = 0$, no superheat is possible, and if $W_{LS} < 0$, a negative superheat is theoretically possible. The latter statement is obviously incorrect, since no amount of catalytic activity by the ebullator can make the bulk-phase transformation proceed spontaneously in a direction which increases the total free energy of the system.

[1] This work was done at Argonne National Laboratory, Lemont, Ill.
[2] Rose Polytechnic Institute.
[3] Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Heat Transfer Division of The American Society of Mechanical Engineers and presented at the Heat Transfer and Fluid Mechanics Institute, Stanford, Calif., June 21–23, 1956.
Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, July 2, 1956.

Sarukhanian (7) assumes that bubble formation occurs at convex elements of surface roughness whose radius is given by the Gibbs equation

$$r^* = \frac{2\sigma_{LG}}{P_G' - P_L} = \frac{2\sigma_{LG}}{\Delta P}\frac{\rho_L}{\rho_L - \rho_G} \quad [2]$$

where P_G' is the vapor pressure of the superheated liquid in the bubble of critical radius, and P_L is the external pressure. Upon applying the Thomson correction, the right-hand equation is obtained, where ΔP is the superheat-vapor-pressure difference, and ρ_L and ρ_G are the average densities.

It is apparent, therefore, that these hypothesized mechanisms for bubble nucleation must be considered as a necessary preliminary to the development of a kinetic theory of ebullition. In this paper the theories of heterogeneous nucleation of Volmer (8, 9), Fisher (10), and Turnbull (11) are extended to include consideration of ebullition nucleation from wetted and unwetted surface projections and cavities. It is shown that the initiation of ebullition almost always, and the continuation of ebullition always proceeds from small quantities of vapor or gas trapped in cavities in the solid surface.

Nucleation at Plane Surfaces

Volmer (8) has calculated the work of nucleating a phase G in the transformation $L \rightarrow G$, on a plane solid surface denoted by S. Let the contact angle between the G-embryo and S be θ as shown in Fig. 1. Then the free energy, or the reversible work required to form an embryo having the shape of a spherical sector is given as

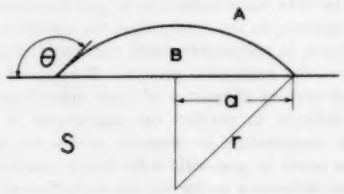


Fig. 1 Embryo on Flat Surface

$$W = \left[\frac{4\pi r^3(\mu_G - \mu_L)}{3v_G} + 4\pi r^2\sigma_{LG}\right] f(\theta) \quad [3]$$

where

$$f(\theta) = \frac{2 + 3\cos\theta - \cos^3\theta}{4} \quad [4]$$

μ is the free energy per mole in the phases of unlimited extent, and v is the molal volume.

For a nucleus of critical size this becomes

$$W_{max} = \frac{16\pi\sigma_{LG}^3 v_G^2}{3(\mu_G - \mu_L)^2} f(\theta) \quad [5]$$

This formula must be used with care since it is derived, in accordance with the method of Gibbs (12), from the PV work and work of forming the interfaces. An exact expression is therefore

$$W_{max} = \frac{16\pi\sigma_{LG}^3}{3(P_G' - P_L)^2} f(\theta) \quad [5a]$$

where P_G' and P_L are the pressures of the G-embryo and the L-phase. For an embryo to be a nucleus, the chemical potentials of the two phases must be equal, or

$$\mu_G' = \mu_L \quad [6]$$

where the prime here refers to the G-phase in the embryo. Then

$$\mu_G' - \mu_G = \int_{P_L}^{P_G'} v_G dP \sim v_G(P_G' - P_L) \quad [7]$$

Hence, Equation [5] is allowable only if the embryonic phase is essentially incompressible, or if the ratio of the pressure within the embryo and the external pressure does not differ too greatly from unity. For our purposes, the use of chemical potentials offers no particular advantages, and the more exact expression, Equation [5a], will be used.

Fisher (10) has applied the methods of nucleation theory (9, 13) to the nucleation of vapor bubbles in liquids under high hydrostatic tension. In this case the pressure within the nucleus is negligible compared to the large negative pressures of the liquid. This theory can be applied to the superheating of liquids directly.

The work of formation of a vapor nucleus of critical size is given by Equation [5a]. For homogeneous nucleation in the absence of a solid surface, $f(\theta) = 1$. The rate of formation of bubbles of vapor in a mole of superheated liquid is then

$$\frac{dn}{dt} = \frac{NkT}{h} \exp\left[-\frac{(\Delta f_0^* + W_{max})}{kT}\right] \quad [8]$$

where N is the Avogadro number and Δf_0^* is the free energy of activation for the motion of an individual molecule of liquid past its neighbors into or away from the bubble surface. Combining Equations [2], [5a], and [8], this becomes

$$\frac{dn}{dt} = \frac{NkT}{h} \exp\left\{-\left[\Delta f_0^* + \frac{16\pi\sigma_{LG}^3}{3(\Delta P)^2}\left(\frac{\rho_L}{\rho_L - \rho_G}\right)^2\right] \Big/ kT\right\} \quad [9]$$

Fisher shows that the error of neglecting Δf_0^* is at most 5 or 10 per cent, and also that the fracture pressure is remarkably insensitive to the waiting time for the first bubble. Based on these two observations, Equation [9] can be solved for the superheat-vapor-pressure difference

$$\Delta P = \frac{\rho_L}{\rho_L - \rho_G}\left(\frac{16\pi\sigma_{LG}^3}{3kT \ln \frac{NkT}{h}}\right)^{1/2} \quad [10]$$

For moderate degrees of superheat, the Clausius-Clapeyron relationship can be employed to give

$$\Delta T = \frac{\rho_L}{(\rho_L - \rho_G)\lambda P}\left(\frac{16\pi RN\sigma_{LG}^3}{3 \ln \frac{RT}{h}}\right)^{1/2} \quad [11]$$

where λ is the molal latent heat of evaporation.

In Equation [10] it will be noted that the temperature has only a moderate influence. Hence ΔP for a pure liquid would be expected to approximate the theoretical fracture pressure, which is in the neighborhood of hundreds to thousands of atmospheres. Experimentally it has been found very difficult to support superheats corresponding to ΔP's in excess of a few atmospheres. Nucleation from the homogeneous liquid therefore can be dismissed from consideration.

Similarly, the superheat-vapor-pressure difference at a plane solid surface can be shown, using Equations [2], [5a], and [8] to be

$$\Delta P = \frac{\rho_L}{\rho_L - \rho_G}\left[\frac{16\pi\sigma_{LG}^3 f(\theta)}{3kT \ln \frac{6N^2/^3 kT}{h}}\right]^{1/3} \quad [12]$$

In Equation [4] it is seen that if $\theta = 180$ deg, $f(\theta) = 0$, and no superheat is possible. Such a contact angle has never been recorded, the largest found experimentally being about 140 deg. Paraffin, which is commonly considered not to be wetted by water, exhibits an average contact angle of about 95 deg if quite smooth. The theoretical ΔP for water in contact with a plane paraffin surface is about 60 per cent of that of the pure liquid, or about 800 atm. Hence nucleation from plane surfaces also can be dismissed from consideration in ebullition.

Frenkel (13) states the work of rupturing the solid-liquid interface, given by Equation [1], should be substituted for the free interfacial energy σ_{LG} in the Volmer theory. This is difficult to justify, since the nucleus is not yet ready to grow spontaneously until it assumes the shape of a spherical section of critical radius of curvature. In any case, the calculated ΔP is reduced only by a factor of 2 if $\theta = 90$ deg, and hence this question is unimportant for ebullition considerations.

It is seen that the hypothesis (6), that no superheat should be possible when $\theta = 90$ deg, must be rejected. This hypothesis results from a consideration of the relative values of σ_{LS} and σ_{SG}, but fails to take into account that the work of separation is still positive at a contact angle of 90 deg.

Nucleation at Curved Surfaces

Nucleation from a surface projection, proposed by Jakob (5) and Sarukhanian (7), is now considered. Suppose the solid projection to be spherical, of radius r_s, and the vapor embryo, of critical radius r^*, to have an included angle, 2β, while the included angle of the gas-solid interface is 2α (Fig. 2). The work forming the vapor phase is

$$W = \sigma_{LG}A_{LG} + (\sigma_{SG} - \sigma_{LS})A_{SG} - (P_G' - P_L)V_G \quad [13]$$

where A and V represent areas and volume, respectively. It can be shown that this reduces to

$$W = \frac{16\pi\sigma_{LG}^3\phi(\alpha, \beta)}{3(\Delta P)^2}\left(\frac{\rho_L}{\rho_L - \rho_G}\right)^2 \quad [14]$$

where

$$\phi(\alpha, \beta) = \frac{3}{2}\sin^2\beta\left(\frac{1}{1+\cos\beta} + \frac{\cos\theta}{1+\cos\alpha}\right) - \frac{\sin^3\beta}{4}\left[\tan^3\frac{\beta}{2} + 3\tan\frac{\beta}{2} - \left(\tan^3\frac{\alpha}{2} + 3\tan\frac{\alpha}{2}\right)\right] \quad [15]$$

and θ, the contact angle $= 180 - (\beta - \alpha)$. The corresponding expression for the superheat-vapor pressure difference, ΔP, may be found by substituting $\phi(\alpha, \beta)$ for $f(\theta)$ in Equation [12]. In general, $\phi(\alpha, \beta)$ is of the same order of magnitude as $f(\theta)$ for a vapor nucleus of critical radius of curvature. This can be shown by considering the special case of a hemispherical projection of critical radius r^* on a flat surface (Sarukhanian's case), with $\theta = 90$ deg (Fig. 3). The last two terms of Equation [13] vanish, so that the work is simply the work of forming the liquid-vapor interface. This yields

$$W = \frac{8\pi\sigma_{LG}^3}{(P_G' - P_L)^2} \quad [16]$$

actually higher than the work of nucleation in the homogeneous phase (Equation [5a], $f(\theta) = 1$) since no PV work is received by the surroundings in this case. If the contact angle is greater than 90 deg, the included angle, and hence the area, of the spherical projection can be reduced; but it is obvious that no fracture pressure or superheat approaching observed values can be obtained in this way. If the radius of the projection is less than critical, less work will be required to fracture the solid-liquid bonds, but the bubble must still grow to the critical radius before it can act as a nucleus. In the limiting case, as the radius of the projection approaches zero, the work of forming a nucleus approaches that of a plane surface, treated previously. If the radius is greater than critical, the work will be greater than given by Equation [15]. Hence, nucleation from surface projections also can be dismissed from consideration.

At first glance, this may seem to be surprising, in view of the obvious analogy to the fracture of supercooled liquids, such as glasses, by nucleation from surface cracks. On second thought, it will be seen that the analogy is inexact. In the previous case nucleation proceeds by fracturing liquid-solid bonds, while nucleation from surface cracks proceeds entirely by fracturing bonds within the homogeneous phase. It is actually more difficult for the liquid to pull away from a surface projection than from the plane surface, and, correspondingly, less difficult for it to pull away from a cavity in the solid surface, since it tends thereby to reduce the area of its boundaries.

Hence if the foregoing analysis is extended to consider surface cavities, it is not unexpected that a more favorable situation for nucleation obtains (Fig. 4). It is now possible to increase the PV work which the surroundings receive to the point where the

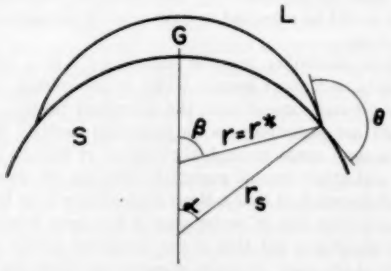


Fig. 2 Nucleation From a Spherical Surface

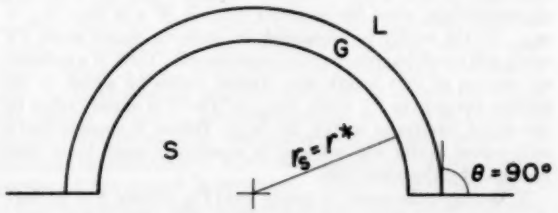


Fig. 3 Nucleation From a Hemispherical Projection

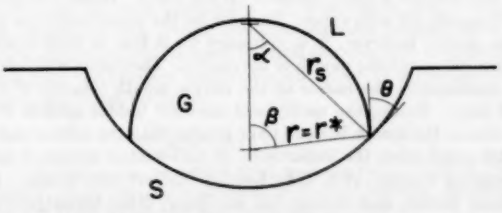


Fig. 4 Nucleation From a Spherical Cavity

net work approaches zero. Equations [14] and [15] can be applied, with the alteration of one sign

$$W = \frac{16\pi\sigma_{LG}^3}{3(\Delta P)^2}\left\{\frac{3}{2}\sin^2\beta\left(\frac{1}{1+\cos\beta}+\frac{\cos\theta}{1+\cos\alpha}\right)\right.$$
$$\left.-\frac{\sin^3\beta}{4}\left[\tan^2\frac{\beta}{2}+3\tan\frac{\beta}{2}+\left(\tan^2\frac{\alpha}{2}+3\tan\frac{\alpha}{2}\right)\right]\right\}$$
$$\left(\frac{\rho_L}{\rho_L-\rho_G}\right)^2 = \frac{16\pi\sigma_{LG}^3}{3(\Delta P)^2}\,\phi'(\alpha,\beta)\left(\frac{\rho_L}{\rho_L-\rho_G}\right)^2 \quad \ldots\ldots [17]$$

where now $\theta = 180 - \alpha - \beta$.

If r^* and θ are fixed, the specification of one included angle will fix the system. It may be observed that in both Equations [15] and [17], ϕ and ϕ' vanish if $\beta = 0$; and it would seem superficially that the work of forming a nucleus can be made arbitrarily small by choosing a small enough cavity or projection. However, this cannot be correct; else no superheat or tensile stress could be sustained by liquids in contact with any solid surfaces. The fallacy lies in the fact that the work of nucleation must include the total work up to the point where the bubble is ready to grow spontaneously. Hence, while it requires arbitrarily small work to tear the liquid away from an arbitrarily small portion of the surface, whether it be a projection, cavity, or a plane area, each case must be considered separately to determine whether nucleation of the vapor bubble will proceed spontaneously from this point. If two conditions are met, the bubble will grow spontaneously on a solid surface:

1 The bubble must be of critical radius; i.e., $r = r^*$.

2 The radius of curvature of the vapor-liquid interface should not become less than the critical as the bubble grows; i.e., $r \geqq r^*$, $V > V^*$. Note that this does not imply that $(dr)/(dV)$ must always be positive. In certain cases, such as necked cavities, or highly nonwetted cavities, the latter point is especially pertinent. Referring to the cases of a flat plane or of a projection, it is seen that the application of these two conditions requires that the work of nucleus formation be given by Equations [5a] and [14].

Returning now to the problem of nucleation in cavities, in the case of a conical cavity with rounded bottom (Fig. 5), the radius of curvature of the meniscus continuously increases as the bubble grows from the apex to a macroscopic size, provided the solid is well wetted. If the bubble is approximately spherical at critical radius, the work of formation will be given by Equation [5a], approximately, since for complete wetting ($\theta = 0$ deg), $\sigma_{SG} = \sigma_{LG}$. If the cavity be elongated, in order to obtain more PV work, not much improvement is experienced. Thus, if a cylindrical section of unit length and critical radius be added to the bubble, the gain in PV work, $(2\sigma_{LG}/r^*)(\pi r^{*2})$ is exactly offset by the added interfacial energy, $2\pi r^*\sigma_{SG}$. Hence, it appears that a well-wetted cavity will not lead to superheats much lower than predicted by Equation [10].

If, however, the cavity is poorly wet (Fig. 6) (say, $\theta = 90$ deg), the situation is entirely different. The radius of curvature of the meniscus begins at a very small value at the apex, but rapidly becomes large, and may approach infinity, or a negative value after it has climbed a short distance up the walls. Hence the cavity will rapidly fill with vapor. In order for the vapor to expand past the cavity, however, it is necessary for it first to form a hemispherical cap at the mouth of the cavity. The radius of this cap is approximately the radius of the cavity mouth (exactly if $\theta = 90$ deg). Hence, the cavity will nucleate bubble growth if its radius at the mouth is equal to or greater than the critical radius. This emphasizes the importance of well-wetted container walls found by Harvey, et al. (14), Kenrick, Gilbert, and Wismer (15), Mead, Romie, and Guibert (4), and many other investigators in making static determinations of maximum superheat or tensile

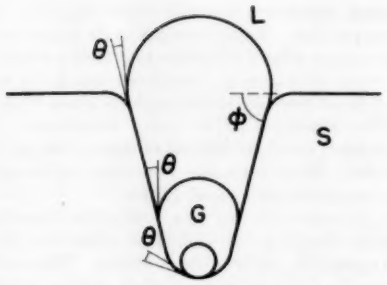


FIG. 5 SUCCESSIVE BUBBLE PROFILES—WETTED CAVITY

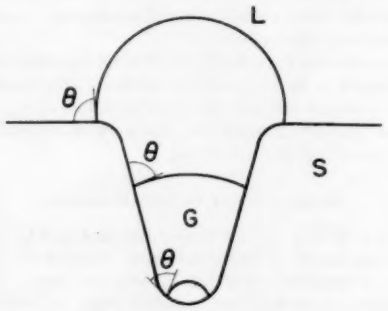


FIG. 6 SUCCESSIVE BUBBLE PROFILES—UNWETTED CAVITY

stress. At the same time, it confirms the necessity of having the walls as smooth as possible, as indicated by the superiority of glass over metallic surfaces (17, 18), in attempting to obtain large values of static tensile stress or superheat. A good review of the extensive literature on this subject is given by Blake (18).

Nucleation in a wedge-shaped crevice may now be discussed. If the walls of the crevice are well wetted, approximately the same considerations will apply as in a cavity, and the theoretical superheats and fracture pressures will be far in excess of experimental values. If the walls are nonwetted, the liquid can be torn away from an arbitrarily small area of apex with very little expenditure of work, but the resulting vapor phase cannot grow until the liquid-solid bond has been fractured along the entire length of the crevice. Hence, nucleation in the absence of a pre-existing vapor phase will normally take place at nonwetted cavities rather than crevices. This distinction is important, since most newly formed metallic surfaces are covered predominantly with grooves rather than cavities, owing to the gouging or cutting action used in producing surface finishes, or in forming the dies or molds which shape the surface. On the other hand, sand-castings; oxidized, such as weathered or anodized, surfaces; or chemically etched surfaces, such as produced by phosphate or chromate treatments, all would be expected to have a predominance of cavities on the surfaces.

One other possibility may be mentioned. If a well-wetted crevice contains a short section which is not wetted, the vapor embryo need not extend past the unwetted portion, and the crevice will act essentially as an unwetted cavity. Such contamination may occur through the deposit of minute quantities of grease and other organic materials from the air, as evidenced by the well-known fact that a clean glass slide will no longer support a continuous film of water once it has been dried in air.

Fisher (10) points out that if the unwetted cavity or crevice has a truly sharp apex, no finite pressure can force the liquid all the way into the depression. Hence it appears probable that, in

almost all cases of initiation of bubble formation from nonwetted surfaces, nucleation proceeds from a pre-existing vapor phase, rather than by tearing the liquid from a portion of the solid boundary. If the liquid is saturated with air, it will have no tendency to rise into the unwetted capillaries, and essentially no superheat or tensile stress can be maintained. If the liquid has been partially or almost completely deaerated, it will tend to rise into the capillaries until the partial pressure of air in the capillary vapor space is in equilibrium with the dissolved air content. Since this is a diffusional process, this may require hours, or even days. If the liquid is deaerated, at one atmosphere pressure, and has negligible vapor pressure at room temperature, the vapor embryo pressure eventually will drop practically to zero, corresponding to a radius of curvature of about 10^{-4} cm in the steep-walled unwetted cavities. The vapor embryos will collapse completely in the wetted cavities; and if the meniscus cannot attain this critical radius of curvature as it advances in an unwetted cavity, the vapor phase also will disappear. If the cavity having a sharp apex is well wetted, the vapor phase will collapse until its radius of curvature is within an order of magnitude of the spacing between molecules in the liquid, so that the cavity is effectively nullified as a nucleation center. Hence, it appears that a large contact angle is essential for the retention of effective vapor-phase nuclei on the solid surface.

If the surface is not wetted, or not wetted in spots, almost certainly there will be present some steep-walled cavities which will retain a vapor phase. If the liquid is now subjected to the slightest increase in temperature or decrease in pressure, the nucleus will begin to grow spontaneously, but cannot leave the cavity unless the mouth of the cavity is equal to or greater than the critical radius, as noted previously. Hence, the radius of the largest unwetted cavity will determine the static superheat or tensile stress; in dynamic phenomena, such as cavitation and ebullition, the superheat or tensile stress is determined by the cavity size-distribution function. Dynamic phenomena are more complex, in that wetted cavities also can function effectively as vapor traps.

If the system is subjected to pre-pressurization at, say, 1000 atmospheres, the radius of curvature of the meniscus in an unwetted cavity with a truly sharp apex would theoretically reach a value of about 10^{-7} cm, if the surface tension did not change with the radius of curvature, as noted by Tolman (19). The definitions of the various terms become imprecise when dealing with such small aggregations of molecules; but there seems to be no reason why the vapor phase should not spontaneously grow again as soon as the pressure is released, providing a concave meniscus (viewed from the liquid phase) of 10^{-6} to 10^{-7} cm can be attained. Hence, one might expect pre-pressurization not to be particularly useful in increasing static superheats or tensile stresses with unwetted surfaces, which was, in fact, observed by Harvey, et al. (14).

If the surface is well wetted, all vapor phases eventually should disappear, if the liquid is not saturated with dissolved gas. Thus, Dean (16) found that large tensile stresses, induced by striking the side of the tube, could be sustained if various surfaces were simply allowed to stand in contact with partially deaerated water for some time. The increase in tensile strength or superheat upon prolonged cavitation (4, 14, 17) indicates that, in most static experiments, the diffusional process controls. Hence, it appears likely that the chief advantage of the pre-pressurization treatment is to increase the driving force for diffusion of the entrapped air into the liquid. Another undeniable advantage of pre-pressurization is that vapor embryos in partially wetted cavities will be collapsed, since a contact angle of at least 90 deg is almost a prerequisite to the attainment of concave menisci of 10^{-6} to 10^{-7} cm radius of curvature. However, such partially wetted cavities will require lower superheats or tensile stresses to initiate nucleation, as given by Equation [17]. As θ increases, $\phi'(\alpha, \beta)$ decreases for constant $r = r^*$. This can be shown more readily for a cylindrical cavity, for which the work of formation of a vapor embryo which fills the cavity is

$$W = 2\pi\sigma_{LG}r^2(1 - \sin\theta) \quad [18]$$

The work of filling the cavity can therefore be made to approach zero as closely as desired by letting θ approach 90 deg. The superheat or tensile stress, in this case, depends on the radius of the mouth of the cavity, as shown in the foregoing.

It is interesting to note that Dean (16) was able to wet his paraffin surfaces by allowing them to remain in contact with partially deaerated water for a day or more, and that thereafter they did not promote bubble formation when the tube was struck a sharp blow. Apparently, the adsorbed air layer was eventually dissolved, after which the paraffin appeared to be wetted. It seems probable, therefore, that bubble nucleation always begins at a pre-existing vapor phase, either as an adsorbed gas layer on a nonwet surface, or as a bubble entrapped in a cavity or groove.

Summary and Conclusions

The nucleation theory of Volmer and Fisher has been extended to the superheating of liquids. It is shown that nucleation of bubbles within the homogeneous of liquid, or at flat or projecting solid surfaces, requires superheat-vapor-pressure differences of the order of magnitude of hundreds of atmospheres, within the range of observed contact angles. This is so much larger than experimentally observed superheats that these possibilities can be dismissed from consideration. Well-wetted cavities also are rejected, since the work of forming a vapor embryo is of the same order of magnitude, irrespective of the depth of the cavity, as the work of forming a critical nucleus of the same radius of curvature in the bulk phase. If, however, the wetted cavities contain air, or other gas, sufficient time must be allowed to allow the gas to dissolve completely. Otherwise, the superheat or tensile stress obtainable will be determined by the radius of curvature of the largest remaining air bubble. Since this is a diffusional process, hours or days may be required; pre-pressurization is useful as a means of accelerating the process. If any nonwetted cavities exist, however, nucleation will preferentially occur there, since it requires only a small amount of work to tear the liquid away from the apex of the cavity. After this the cavity rapidly and spontaneously fills with vapor, so that the superheat is determined, in this event, by the radius of curvature of the mouth of the unwetted cavity. If the unwetted cavity has a truly sharp apex, no finite pressure can force the liquid to occupy the cavity completely. However, pre-pressurization is useful in filling up partially wetted cavities. Nucleation will require less work in unwetted cavities than in unwetted grooves, unless only a small portion of the groove is unwetted. These observations apply equally well to static determinations of superheat, tensile strength, or supersaturation. In the dynamic phenomena of ebullition, cavitation, and effervescence, either wetted or unwetted cavities may contain a vapor or gas phase; and nucleation always occurs preferentially at these boundaries. It is probable that even in static determinations, unless the most rigorous precautions are taken to achieve complete wetting of all cavities, and to dissolve all gas from the cavities, nucleation will occur at a pre-existing vapor or gas-phase boundary.

BIBLIOGRAPHY

1 "Heat Transfer Associated With Nucleate Boiling," by W. M. Rohsenow, Heat Transfer and Fluid Mechanics Institute, Stanford University Press, Stanford, Calif., 1953, pp. 123–141.

2 "Surface Variables in Nucleate Boiling," by C. Corty and A. S.

Foust, Chemical Engineering Progress Symposium Series, No. 17, "Heat Transfer," 1955, pp. 1–12 and 12a.

3 "A Study of the Mechanism of Boiling Heat Transfer," by M. E. Ellion, JPL Memo 20-88, 1954.

4 "Liquid Superheat and Boiling Heat Transfer," by B. R. Mead, F. E. Romie, and A. G. Guibert, Heat Transfer and Fluid Mechanics Institute, Stanford University Press, Stanford, Calif., 1951, pp. 209–216.

5 "Heat Transfer," by Max Jakob, John Wiley & Sons, Inc., New York, N. Y., vol. 1, 1949, p. 624.

6 "Factors That Influence Heat Transfer in Boiling," by R. F. Larson, Heat Transfer and Fluid Mechanics Institute, Stanford University Press, Stanford, Calif., 1953, pp. 163–172.

7 "Wärmeübergang bei Verdampfung," by G. Sarukhanian, *Chemie-Ingenieur-Technik*, vol. 25, 1953, pp. 477–480.

8 "Über Keimbildung und Keimwirkung als Spezialfälle der heterogenen Katalyse," by M. Volmer, *Zeitschrift für Elektrochemie*, vol. 35, 1929, pp. 555–561.

9 "Kinetik der Phasenbildung," by M. Volmer, Steinkopf, Dresden and Leipzig, Germany, 1939.

10 "The Fracture of Liquids," by J. C. Fisher, *Journal of Applied Physics*, vol. 19, 1948, pp. 1062–1067.

11 "Kinetics of Heterogeneous Nucleation," by D. Turnbull, *Journal of Chemical Physics*, vol. 18, 1950, pp. 198–203.

12 "Collected Works," by J. W. Gibbs, Yale University Press, New Haven, Conn., 1948, p. 254.

13 "Kinetic Theory of Liquids," by J. Frenkel, Clarendon Press, Oxford, England, 1946.

14 "On Cavity Formation in Water," by E. N. Harvey, W. D. McElroy, and A. H. Whiteley, *Journal of Applied Physics*, vol. 18, 1947, pp. 162–172.

15 "The Superheating of Liquids," by F. B. Kenrick, C. S. Gilbert, and K. L. Wismer, *Journal of Physical Chemistry*, vol. 28, 1924, pp. 1297–1307.

16 "The Formation of Bubbles," by R. B. Dean, *Journal of Applied Physics*, vol. 15, 1944, pp. 446–451.

17 "The Behavior of Water Under Hydrostatic Tension," by H. N. V. Temperley and L. G. Chambers, Proceedings of the Physical Society of London, England, part I, vol. 58, 1946, pp. 420–436; part II, vol. 58, 1946, pp. 436–443; part III, vol. 59, 1947, pp. 199–208.

18 "The Tensile Strength of Liquids," by F. G. Blake, Jr., NR-014-903 TM No. 9, June 11, 1949.

19 "The Effect of Droplet Size on Surface Tension," by R. C. Tolman, *Journal of Chemical Physics*, vol. 17, 1949, pp. 333–337.

Natural-Convection Heat Transfer From a Horizontal Cylinder Rotating in Air

By DAVID DROPKIN[1] and ARIEH CARMI[2]

This paper presents the results of experimental investigation of convection heat transfer from horizontal rotating cylinders to ambient air. Two hollow copper cylinders were used, one 4.5 in. and another 3.25 in. outside diameter. These cylinders were nickel plated to minimize the heat transfer by radiation and were internally heated by electric cartridge heaters. The surface temperatures were determined by thermocouples peened to the inside surface. The experimental procedure was to investigate systematically each dimensionless parameter which was previously determined by dimensional analysis. The results indicate that, up to a certain value of Reynolds number, rotation has no effect on the heat-transfer coefficient. Above this critical value the heat-transfer coefficient increases as the speed of rotation increases. In the region where the data overlap, the experimental results reported in this paper are in close agreement with those of other investigators. The flow pattern around the cylinder was also investigated. Titanium tetrachloride smoke was used for this study.

Nomenclature

The following nomenclature is used in the paper:

c_p = specific heat of air at constant pressure, Btu/(lb)(F)
C, C_1, C_2, C_3, C_4 = constants
D = outside diameter of test cylinder, ft
g = acceleration of gravity, 4.17×10^8 ft/(hr)^2
Gr = Grashof number, dimensionless, $D^3\rho^2 g\Delta\beta/\mu^2$
h = convective heat-transfer coefficient, Btu/(hr)(sq ft)(F)
k = thermal conductivity, Btu/(hr)(ft)(F)
m = exponent
Nu = Nusselt number, dimensionless, hD/k
Pr = Prandtl number, dimensionless, $c_p\mu/k$
Re = Reynolds number, dimensionless, $DV\rho/\mu$
t_a = bulk temperature of air, deg F
t_f = film temperature = $(t_s + t_a)/2$, deg F
t_s = temperature of test cylinder surface, deg F
V = test cylinder surface velocity, fph
x, y, z = exponents
β = coefficient of volumetric expansion, 1/F
Δ = temperature difference between cylinder surface temperature and bulk air temperature, $t_s - t_a$, deg F
μ = absolute viscosity, lb/(ft)(hr)
ρ = air density, pcf
ϕ = function

Subscript a indicates that the property was determined at bulk air temperature.
Subscript f indicates that the property was determined at film temperature.

Introduction

The research project reported in this paper was started in September, 1952 (1).[3] At that time there were no reliable mathematical relationships from which heat-transfer coefficients by convection, for rotating cylinders, could be determined. Since then two papers have been presented on this subject; one by Anderson and Saunders (2) and the other by Etemad (3). These two papers covered a range of Reynolds numbers from 0 to 65,400. The present paper greatly extends this range; the data presented here cover a range of Reynolds numbers from 0 to 433,000.

The main purpose of this investigation was to obtain data and, if possible, develop a relationship from which heat-transfer coefficients for rotating cylinders could be predicted. The best approach to the solution of this problem seemed to be dimensional analysis in conjunction with experimentation.

Procedure

It appeared that factors which are generally assumed to influence free and forced convection should be included in the new relationship. In that case

$$h = \phi(D, V, \rho, \mu, \Delta, \beta, c_p, k, g) \quad [1]$$

and without much difficulty we may derive the relationship

$$\frac{hD}{k} = C\left(\frac{D^3\rho^2 g}{\mu^2}\right)^x \left(\frac{V^2}{Dg}\right)^y \left(\Delta\beta\right)^z \left(\frac{\mu c_p}{k}\right)^m \quad [2]$$

For the range of air temperatures investigated, the Prandtl number is practically constant and may be included in the constant, thus simplifying the equation as

$$\frac{hD}{k} = C_1\left(\frac{D^3\rho^2 g}{\mu^2}\right)^x \left(\frac{V^2}{Dg}\right)^y \left(\Delta\beta\right)^z \quad [3]$$

To determine the constant C_1 and the exponents x, y, and z the following procedure was used:

Two of the parameters on the right-hand side of Equation [3] were held constant while the third was varied and its effect on the Nusselt number determined. This procedure was repeated for each one of the three dimensionless groups. The relationships actually used were

$$\text{Nu} = C_2\left(\frac{D^3\rho^2 g}{\mu^2}\right)^x \quad [4]$$

$$\text{Nu} = C_3\left(\frac{V^2}{Dg}\right)^y \quad [5]$$

$$\text{Nu} = C_4(\Delta\beta)^z \quad [6]$$

---

[1] Associate Professor of Mechanical Engineering, Department of Thermal Engineering, Cornell University, Ithaca, N. Y.
[2] Automation Engineering Department, Corning Glass Works, Corning, N. Y.

Contributed by the Heat Transfer Division of The American Society of Mechanical Engineers and presented at the Heat Transfer and Fluid Mechanics Institute, Stanford University, Calif., June 21–23, 1956.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, May 11, 1956.

[3] Numbers in parentheses refer to the Bibliography at the end of the paper.

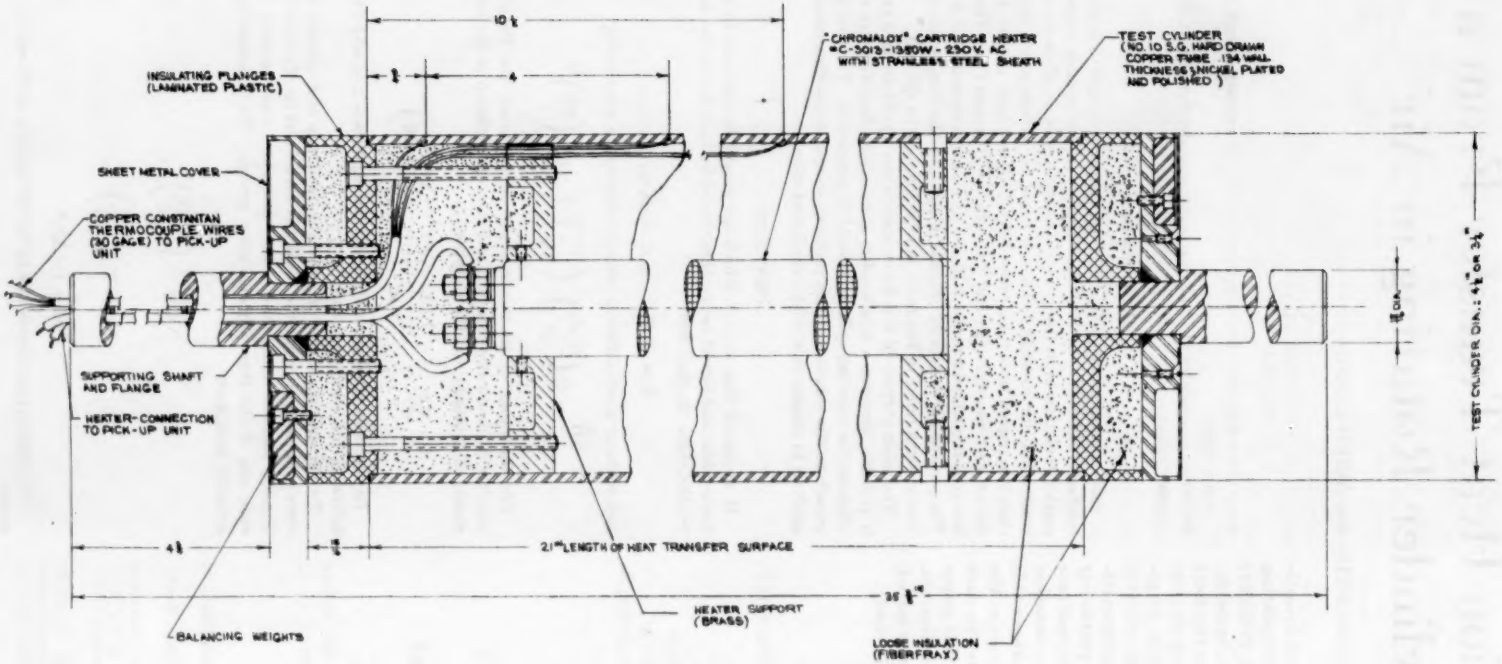


FIG. 1 TEST CYLINDER ASSEMBLY

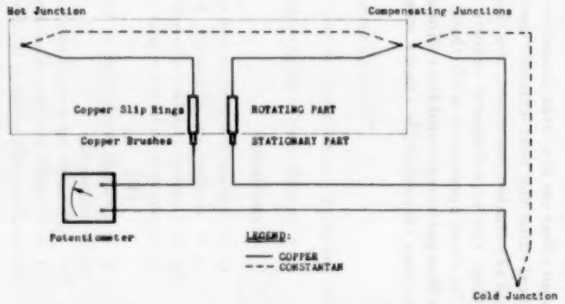


FIG. 2 TYPICAL SCHEMATIC DIAGRAM OF THERMOCOUPLE CIRCUIT

The exponents x, y, and z were then determined from the plots of log Nu versus log $(D^3\rho^2 g/\mu^2)$, log Nu versus log (V^2/Dg), and log Nu versus log $(\Delta\beta)$, respectively. The constant C_1 was evaluated from Equation [3] after the exponents were established.

For Equations [4] and [6], the parameters were varied by varying the air and cylinder surface temperatures while the cylinder diameter and rotational velocity were kept constant. The temperatures, however, had to be so selected that when $(D^3\rho^2 g/\mu^2)$ was varied $(\Delta\beta)$ remained constant, and when $(\Delta\beta)$ was varied $(D^3\rho^2 g/\mu^2)$ was constant.

For Equation [5] the parameter was varied by varying the rotational speed of the cylinder, while both the cylinder surface temperature and air temperature were kept constant.

The experiments were arranged carefully so that sufficient data were taken to evaluate Equations [4], [5], and [6] at both bulk air temperatures and film temperatures.

All data were taken when steady-state conditions were established. It was assumed that steady state was attained when two consecutive readings taken at 1-hr intervals gave the same results.

For each run, the temperature of the surface of the cylinder, the ambient temperature, the rotational speed, and the electrical input to the cylinder were measured.

The following experimental range was covered:
Reynolds number from 0 to 433,000
Nusselt number from 19.1 to 591
Grashof number from $9.5(10)^3$ to $2.27(10)^7$
Cylinder surface temperature from 91.0 to 325.6 F
Ambient temperature from 52.5 to 110.0 F
Rotational speed from 0 to 10,150 rpm
Temperature difference between cylinder surface and ambient air from 6.3 to 230.9 F

Apparatus

The test apparatus consisted of two hollow copper cylinders 21 in. in length and 4.5 in. OD and 3.25 in. OD, respectively. The larger diameter cylinder was used for most test runs, while the smaller cylinder was used mainly to check the effect of cylinder diameter on the heat-transfer coefficient. Both cylinders had their outer surface nickel plated and polished to minimize heat transfer by radiation. The emissivity of each surface was carefully determined and found to be equal to 0.07. As shown in Fig. 1, the ends of the cylinders were insulated to reduce the end heat losses.

A cartridge-type electric heater, placed at the center of the cylinder, supplied the required heat to the cylinder. The electrical input was measured by a voltmeter and ammeter.

The surface temperatures of the test cylinders were measured by three copper-constantan thermocouples. These couples were placed $^3/_4$ in., 4 in., and $10^1/_2$ in., measured from one end of the test cylinder, respectively. The couples were attached to the surface by peening them to the inside surface of the cylinder.

The test cylinder was rotated by an adjustable-speed, electric-motor drive, which in combination with pulleys was capable of producing cylinder speeds up to 12,000 rpm.

Temperature-Measuring System

A great deal of time was devoted to the evaluation of different possible systems that could be used in measuring the surface temperature of the rotating cylinder (4, 5). The study led to the conclusion that a copper-constantan thermocouple system when properly designed should give the best results.

This system, as schematically indicated in Fig. 2, consists of copper slip rings and copper brushes in both sides of the thermocouple circuit. As shown on the diagram and in Figs. 3 and 4, the junctions of the interrupted constantan line were placed in a uniform-temperature chamber so that the rotating junction and the stationary junction were almost touching each other. This method exposed the opposing junctions to the same ambient temperature and thus no net electromotive force was developed due to these junctions. To minimize overheating and scoring, the copper brushes were brought into contact with the copper rings only when measurements were taken. At all other times the brushes were disengaged. The slip rings also were lubricated lightly to reduce brush and slip-ring wear further.

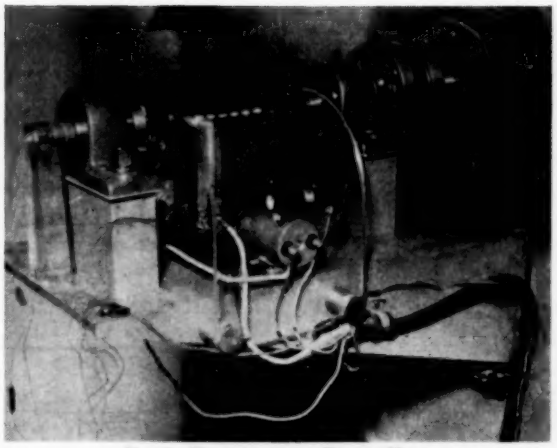


Fig. 3 Thermocouple Pickup Unit

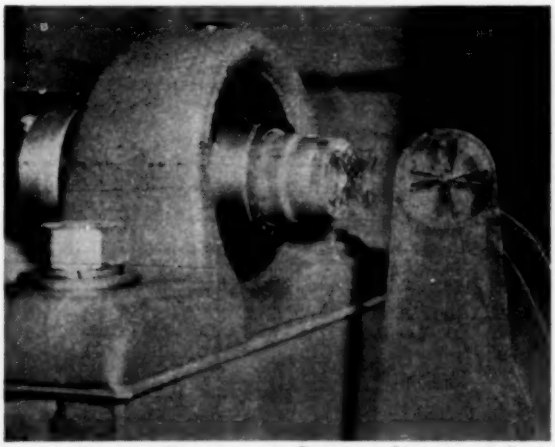


Fig. 4 Uniform Temperature Chamber Opened to Show Compensating Junctions

All thermocouples were connected to precision potentiometers to indicate the temperatures at the measuring junctions.

Other Precautions

All rotating parts of the experimental apparatus were dynamically balanced to minimize radial movement of the test cylinder. To eliminate the influence of stray air currents due to the rotating couplings, two guard plates, one at each end of the cylinder, were provided. Special care was taken to assure that no forced-convection air currents were present within the test room especially in the vicinity of the apparatus (6).

To assure that proper air temperatures were determined, the thermocouples were placed some distance away from the cylinder but at the same height as the cylinder.

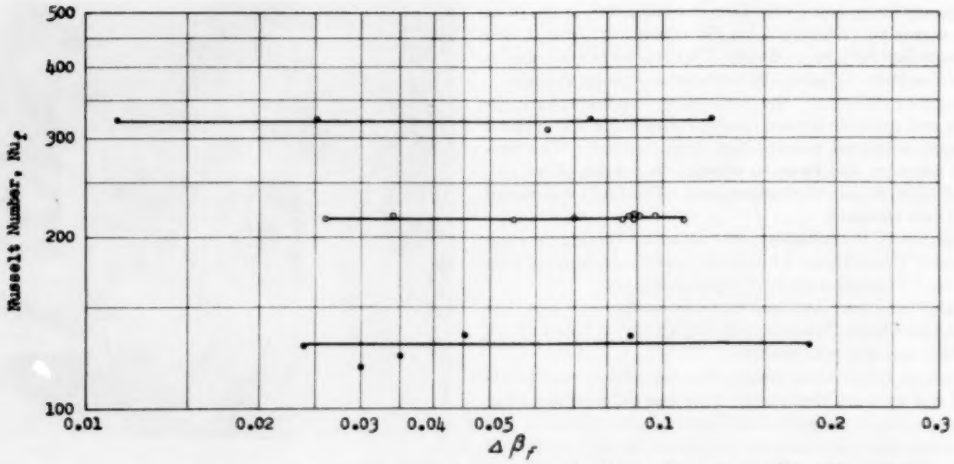


FIG. 5 NUSSELT NUMBER VERSUS $\Delta\beta_f$ FOR FILM PROPERTIES AT CONSTANT $D^3\rho_f^2 g/\mu_f^2$ AND V^2/Dg

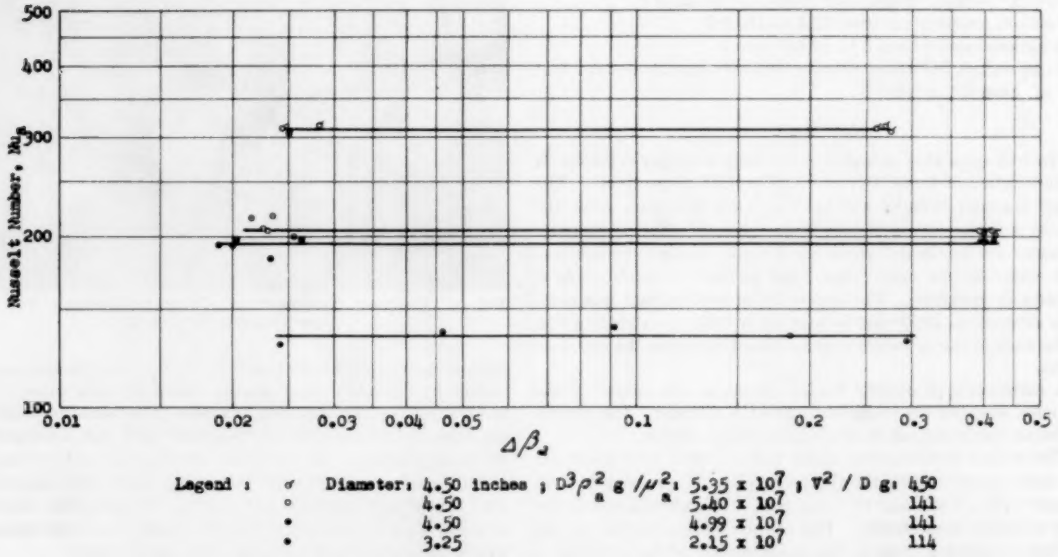


FIG. 6 NUSSELT NUMBER VERSUS $\Delta\beta a$ FOR BULK AIR PROPERTIES AT CONSTANT $D^3\rho_a^2 g/\mu_a^2$ AND V^2/Dg

RESULTS AND DISCUSSION

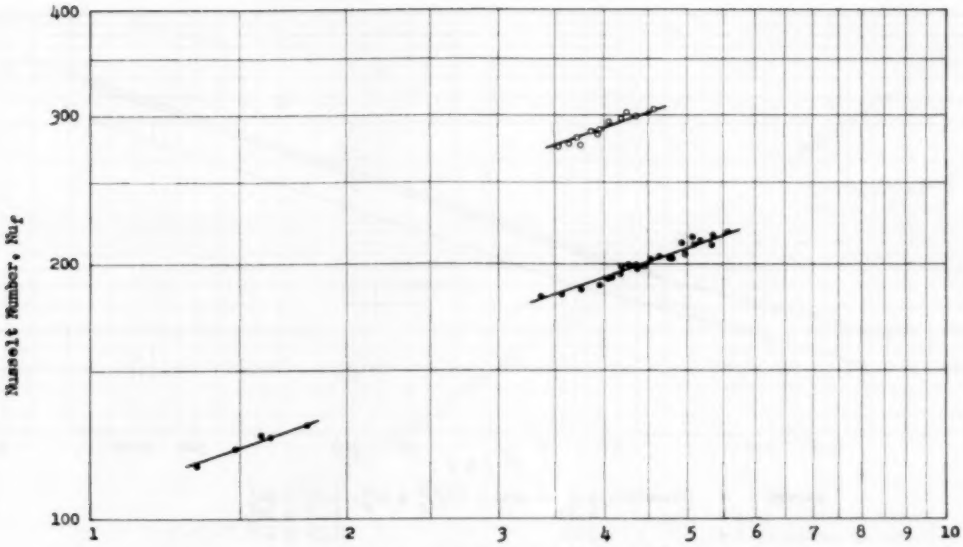
The results of tests on the rotating cylinders are shown in Figs. 5 through 15.

Figs. 5 and 6 give the plots of Nusselt number against the product of $\Delta\beta$, with $(D^3\rho^2 g)/\mu^2$ and V^2/Dg kept constant. For each different constant value of $(D^3\rho^2 g)/\mu^2$ and V^2/Dg there results a different value of Nusselt number. The family of horizontal, parallel lines thus obtained indicate that $\Delta\beta$ has no effect on Nusselt number. This is shown to be true when the physical properties are obtained by using the film temperatures as well as the bulk-air temperatures.

Figs. 7 and 8 are plots of Nusselt number versus $(D^3\rho^2 g)/\mu^2$, with V^2/Dg and $\Delta\beta$ kept constant. A family of parallel, straight lines is indicated when plots are made for different constant values of V^2/Dg and $\Delta\beta$. The slope of these lines is equal to 0.35 and the relationship between Nusselt's number and the parameter $(D^3\rho^2 g)/\mu^2$, for both film and bulk temperatures, may be expressed by equation

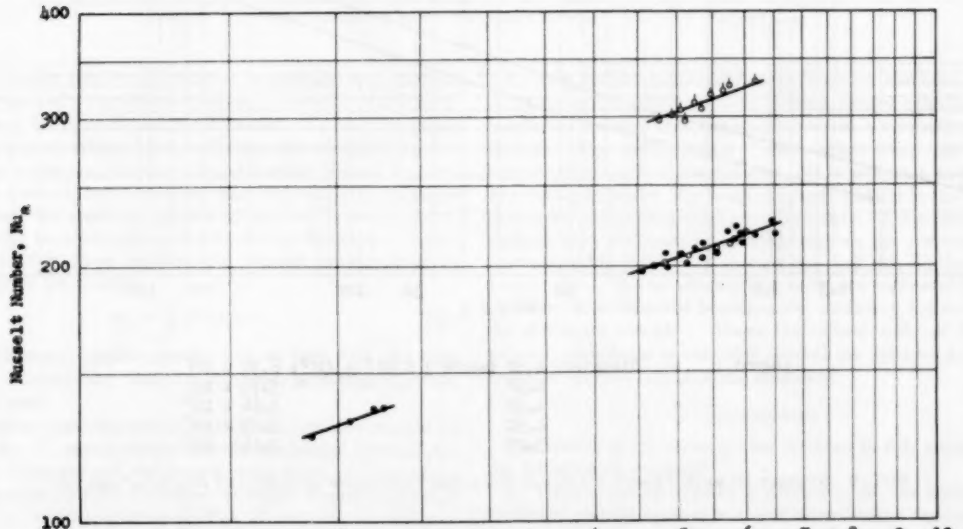
$$Nu = C_3(D^3\rho^2 g/\mu^2)^{0.35} \quad [7]$$

Nusselt number is plotted against the parameter V^2/Dg, with the other parameters kept constant, in Figs. 9 and 10. The resulting curves indicate that the slope of the lines is not constant



Legend :		Diameter:		V^2/Dg:		$\Delta\beta_f$:	
	○	4.50	inches ;	450 ;		0.168	
	•	4.50		141		0.168	
	•	4.50		141		0.0815	
	•	4.50		141		0.084	
	•	3.25		114		0.180	

FIG. 7 NUSSELT NUMBER VERSUS $D^3\rho_f^2 g/\mu_f^2$ FOR FILM PROPERTIES AT CONSTANT V^2/Dg AND $\Delta\beta_f$



Legend :		Diameter:		V^2/Dg:		$\Delta\beta_a$:	
	◊	4.50	inches ;	450 ;		0.184	
	•	4.50		141		0.088	
	•	4.50		141		0.188	
	○	4.50		141		0.085	
	•	3.25		114		0.197	

FIG. 8 NUSSELT NUMBER VERSUS $D^3\rho_a^2 g/\mu_a^2$ FOR BULK AIR PROPERTIES AT CONSTANT V^2/Dg AND $\Delta\beta_a$

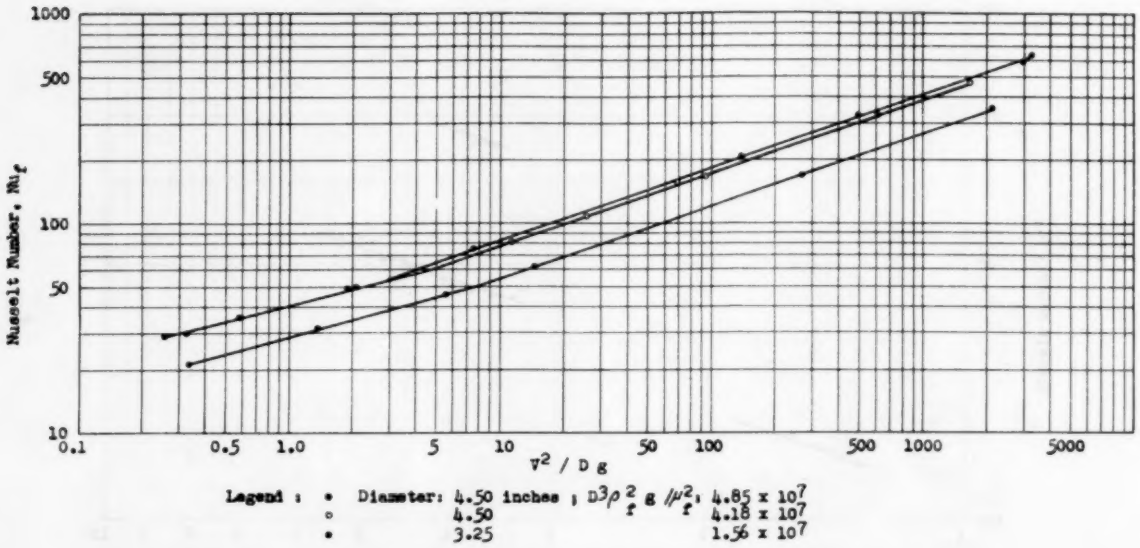


FIG. 9 NUSSELT NUMBER VERSUS V^2/Dg FOR FILM PROPERTIES AT CONSTANT $D^3\rho_f^2 g/\mu_f^2$ AND $\Delta\beta_f$

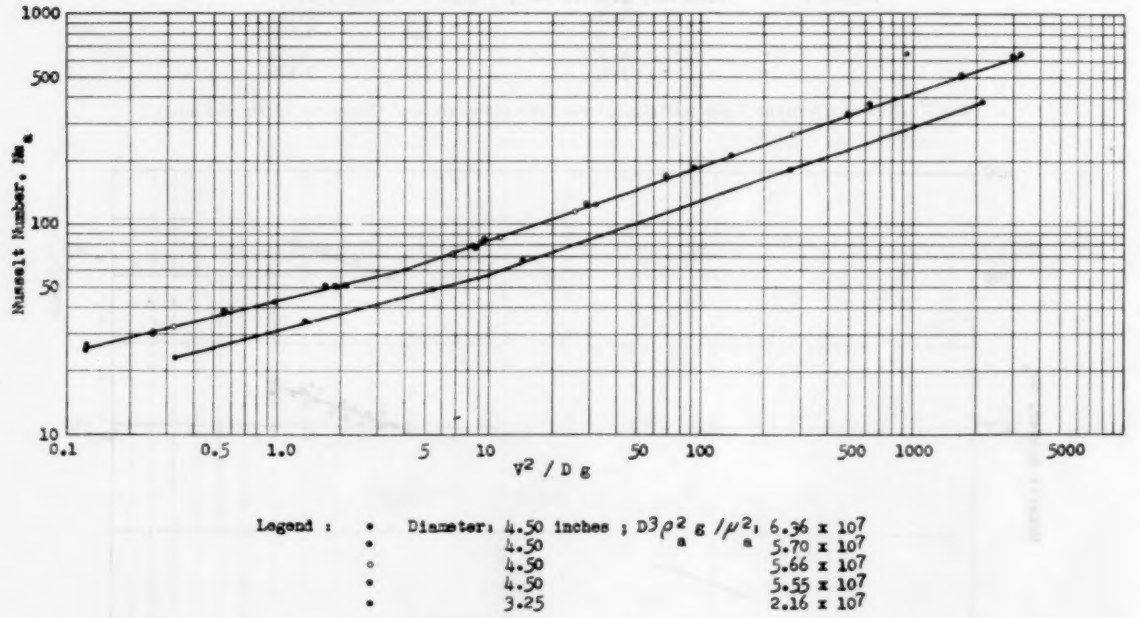


FIG. 10 NUSSELT NUMBER VERSUS V^2/Dg FOR BULK AIR PROPERTIES AT CONSTANT $D^3\rho_a^2 g/\mu_a^2$ AND $\Delta\beta_a$

throughout the range. It increases as the rotational speed increases, but at a certain critical value of V^2/Dg the slope stabilizes and becomes constant and equal to 0.35.

For the major portion of the curves, and for both film and bulk properties, the relationship between Nusselt number and V^2/Dg may be expressed by

$$\mathrm{Nu} = C_3(V^2/Dg)^{0.35} \quad [8]$$

Thus for rotational speeds above the critical V^2/Dg, Equation [3] may be rewritten as

$$\frac{hD}{k} = C_1\left(\frac{D^3\rho^2 g}{\mu^2}\right)^{0.35}\left(\frac{V^2}{Dg}\right)^{0.35}$$

and combining the two parameters on the right-hand side of the foregoing equation, we get

$$\mathrm{Nu} = C_1(\mathrm{Re})^{0.7} \quad [9]$$

Nusselt number is plotted against Reynolds number in Figs. 11 and 12. These curves indicate that up to a certain value of

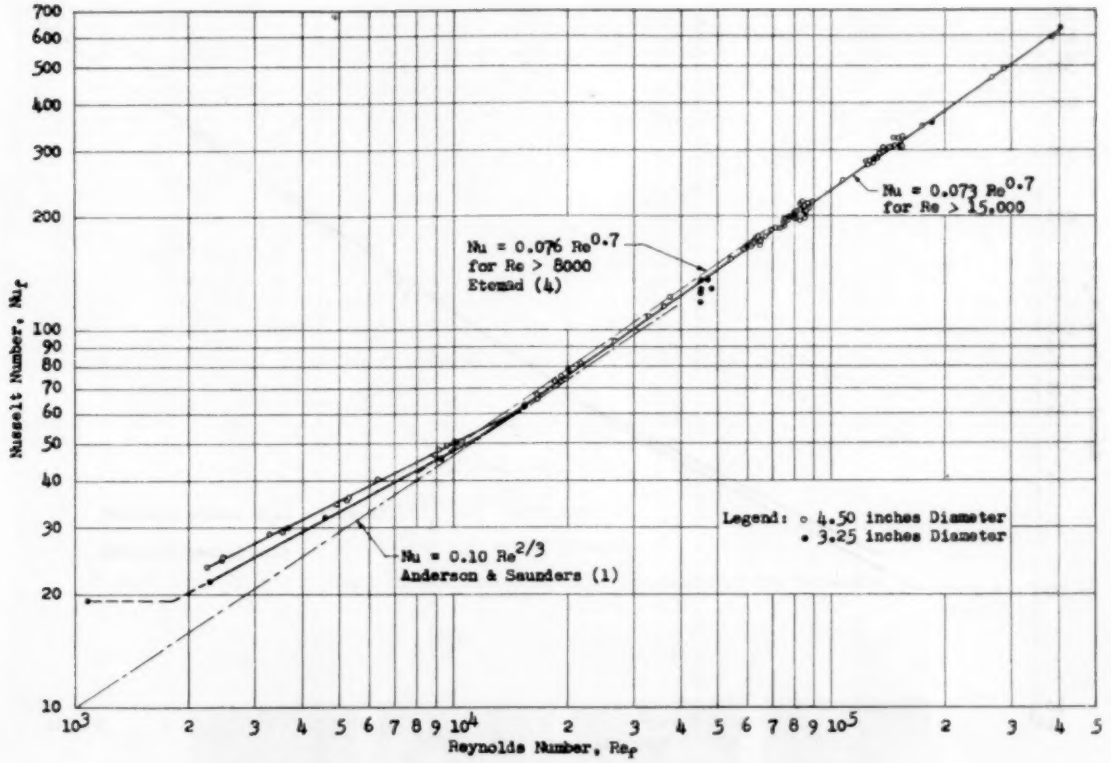


FIG. 11 NUSSELT NUMBER VERSUS REYNOLDS NUMBER FOR FILM PROPERTIES

Reynolds number the Nusselt number is constant and approximates the value of a stationary cylinder. As the Reynolds number increases the Nusselt number increases. Up to a Reynolds number of approximately 15,000, both free convection and the convection due to rotation influence the value of the Nusselt number. Above this number, free convection becomes relatively unimportant and only the speed of rotation affects the Nusselt number. For Reynolds numbers above 15,000, within the range of temperatures used in these experiments, Nusselt number may be computed from the equation

$$Nu = 0.073(Re)^{0.7} \quad [10]$$

This relationship applies equally well to both film and bulk properties. Consistency must be observed, however, when the equation is used.

In the region where the curves overlap, the present curve, as indicated in Fig. 11, agrees closely with the theoretical curve of Anderson and Saunders and the experimental curve of Etemad.

A straight-line plot for the region in which Nusselt number is influenced by rotation is given in Figs. 13 and 14. In this plot Nusselt number is drawn against the empirical parameter $(0.5\ Re^2 + Gr)$. It should be noted that this parameter includes the effect of rotation as well as free convection. For both film and bulk properties, Nusselt number may be computed from

$$Nu = 0.095(0.5\ Re^2 + Gr)^{0.38} \quad [11]$$

The region of Reynolds numbers between 0 and 2500 was investigated by the use of titanium tetrachloride. Fig. 15 shows the smoke pattern for four different Reynolds numbers. In Fig. 15(*a*) the smoke pattern of a stationary, heated cylinder is indicated. It is noted that the smoke rises in a normally expected manner. Fig. 15(*b*, *c*, and *d*) were taken when the cylinder was rotating in a clockwise direction. It is seen, that, as the speed of rotation increases, the breakaway point shifts in the direction of rotation until a tangential point is reached. These photographs indicate that the rotating cylinder opposes the free-convection currents on the downward moving side and aids on the upward moving side. The net effect of this, in the low regions of Reynolds numbers, is to keep the heat-transfer coefficient the same as for the stationary cylinder. Above the critical value of Reynolds number, turbulence is indicated around the cylinder and an increase in rotation increases the coefficient.

CONCLUSIONS

The results of the investigation reported in this paper lead to the following conclusions:

1 When a heated cylinder is rotated in air, the heat-transfer coefficient may be located in one of three distinct regions: (*a*) in a region where its value does not change with speed of rotation; (*b*) in a region where its value increases with the speed of rotation and is influenced by both speed and free convection; and (*c*) a region where its value is proportional only to the speed of rotation.

2 Within the limits of temperatures used in the tests and for Reynolds numbers larger than 15,000, the heat-transfer coefficients may be computed from

$$Nu = 0.073(Re)^{0.7}$$

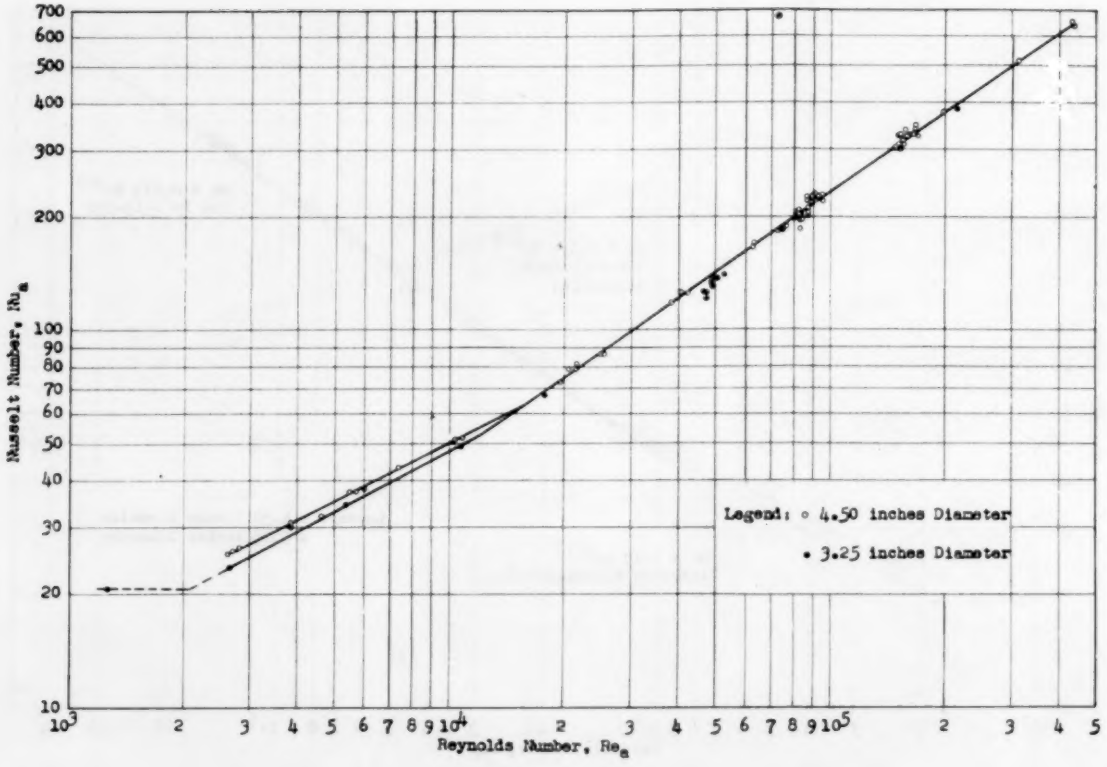


FIG. 12 NUSSELT NUMBER VERSUS REYNOLDS NUMBER FOR BULK AIR PROPERTIES

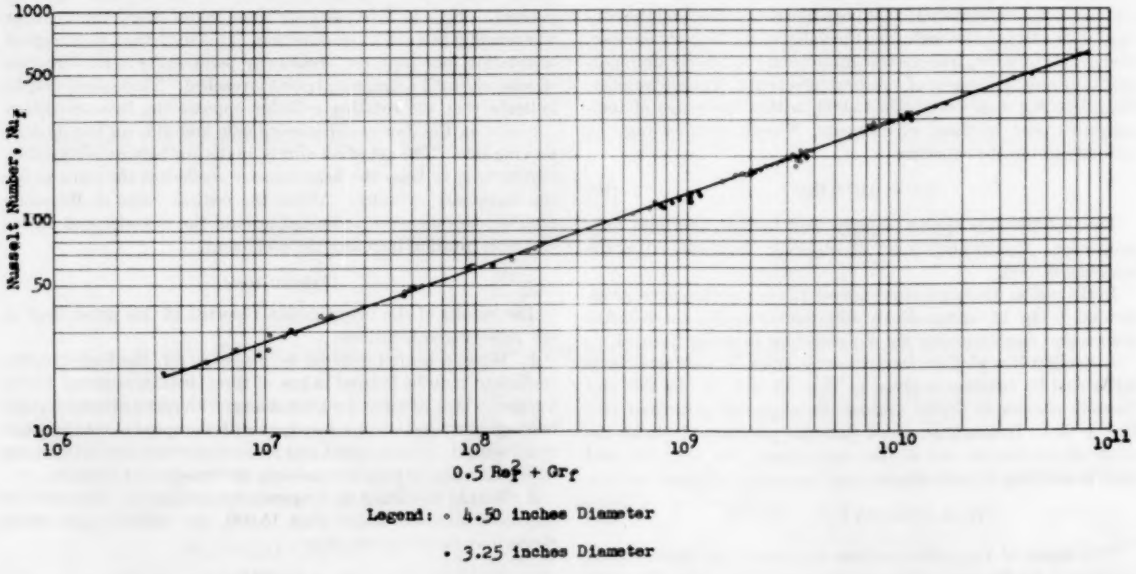


FIG. 13 NUSSELT NUMBER VERSUS $(0.5\ Re_f^2 + Gr_f)$ FOR FILM PROPERTIES

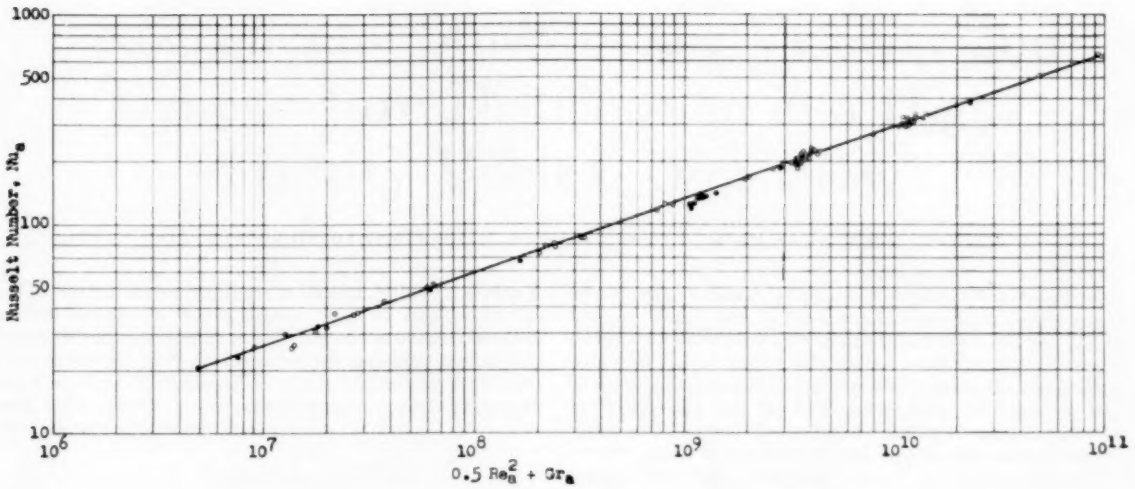


FIG. 14 NUSSELT NUMBER VERSUS ($0.5\ Re_a^2 + Gr_a$) FOR BULK AIR PROPERTIES

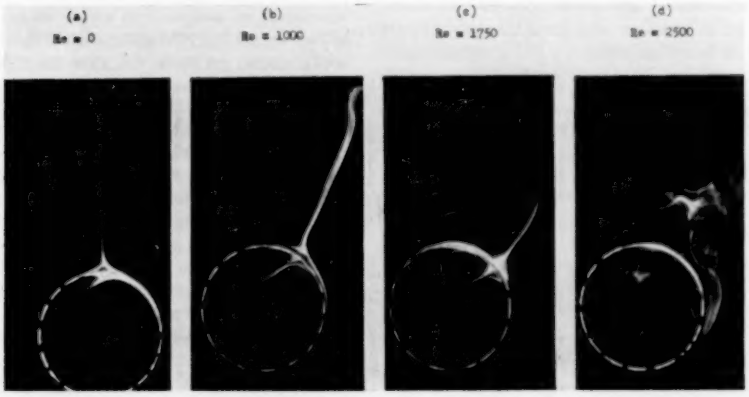


FIG. 15 AIR-FLOW PATTERNS AROUND TEST CYLINDER AT VARIOUS REYNOLDS NUMBERS

3 The entire region in which the heat-transfer coefficient is influenced by rotation may be expressed by equation

$$Nu = 0.095(0.5\ Re^2 + Gr)^{0.48}$$

BIBLIOGRAPHY

1 "The Dissipation of Heat from Rotating Cylinders: Experimental Apparatus," by Arieh Carmi, MS thesis, Cornell University, Ithaca, N. Y., June, 1953.

2 "Convection from an Isolated Heated Horizontal Cylinder Rotating About Its Axis," by J. T. Anderson and O. A. Saunders, Proceedings of the Royal Society of London, England, series A, vol. 217, 1953, p. 555.

3 "Free-Convection Heat Transfer from a Rotating Horizontal Cylinder to Ambient Air With Interferometric Study of Flow," by G. A. Etemad, Trans. ASME, vol. 77, 1955, pp. 1283–1289.

4 "Methods of Connection to Revolving Thermocouples," by P. R. Tarr, NACA RM E50J23a, January, 1951.

5 "Investigation of the Transmission of Thermoelectric Voltages with Sliding Contacts," by D. Dropkin and C. E. Jones, unpublished research, Cornell University, Ithaca, N. Y.

6 "Heat Transfer by Convection from a Horizontal Cylinder Rotating in 'Still' Air," by Arieh Carmi, PhD thesis, Cornell University, Ithaca, N. Y., June, 1955.

Experimental Study of the Velocity and Temperature Distribution in a High-Velocity Vortex-Type Flow[1]

By J. P. HARTNETT[2] and E. R. G. ECKERT,[3] MINNEAPOLIS, MINN.

The vortex tube represents a simple device in which a particular type of vortex motion may be studied in the laboratory in an attempt to obtain a better undertstanding of such flows. Such an investigation has been pursued in the Heat Transfer Laboratory of the University of Minnesota. The present paper summarizes the major results of this vortex-tube investigation.

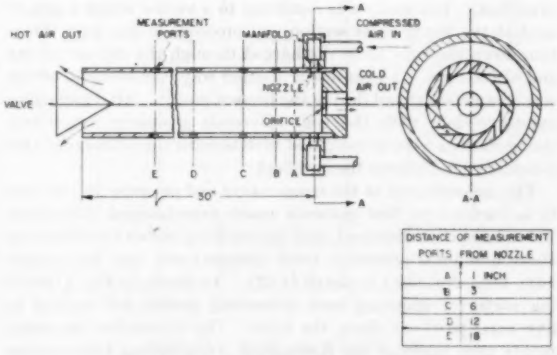
Introduction

IT IS well known that a high-velocity gas flow over a flat plate separates the flow into a region of low total temperature near the surface and a high total temperature out toward the free stream. Other rectilinear motions also demonstrate such a separation of the flow into regions of low and high total temperatures. Such an energy separation has been found to be much more pronounced in vortex motion than in rectilinear flows. Evidence of the marked separating effect of vortex flow is found in experiments showing very low recovery temperature on the rear portion of a cylinder in a high-velocity air flow normal to the cylinder axis (3, 5).[4]

Another device which vividly demonstrates this separation effect of vortex flow is the so-called vortex tube, first patented by George Ranque in 1931, and introduced into the United States in 1945. This device consists of a simple tube, such as shown in Fig. 1, in which compressed air is introduced and expanded in nozzles which are directed tangentially to the tube surface. This air is then discharged through openings on both ends of the tube; the opening at the nozzle cross section removes the air from the central region, while at the far end of the tube the air is discharged from the outer layers. The air leaving the central orifice is found to be at a considerably lower total temperature than the air entering the tube, while the air leaving the far end of the tube is at a higher total temperature than the inlet air.

Test Apparatus and Procedure

The primary consideration in the design of the vortex tube was the establishment of a well-defined vortex in a tube of sufficient diameter to allow the insertion of probes without causing a major disturbance of the flow. In addition, it was felt that the tube should be transparent to allow flow-visualization studies. As a consequence, a 3-in-diam plexiglas tube of 30-in. length was selected, Figs. 1 and 2. The manifold section, into which the compressed air flows, is fabricated of transparent plastic. Eight nozzles of $^3/_{8}$-in. diam equally spaced around the circumference of the tube and rounded at their entrance direct this compressed air tangentially into the main vortex tube. Adjacent to this nozzle cross section the so-called "cold-end" orifice section is located and is so constructed that various size orifices can be inserted from 0 in. (completely closed) up to 1 in. diam. For the experiments described in this paper the "cold end" was completely closed by a diaphragm, as shown in Fig. 2. At the other end of the 30-in-long tube a 60-deg cone-shaped valve is located. This geometry was chosen to preserve flow symmetry, and with the noted exceptions this is the valve used throughout the test program reported herein. The only structural parts penetrating the air stream as it leaves the vortex tube are the rectangular cross sections which are part of the valve holder. In all tests the valve was opened to a position indicated in Fig. 2.



DISTANCE OF MEASUREMENT PORTS FROM NOZZLE	
A	1 INCH
B	3
C	6
D	12
E	18

Fig. 1 Three-In-Diam Vortex Tube

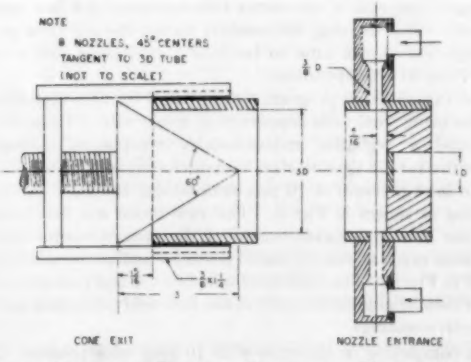


Fig. 2 Details of Exit and Entrance of Vortex Tube

[1] Publication of the Heat Transfer Laboratory, University of Minnesota, Minneapolis, Minn.

[2] Associate Professor of Mechanical Engineering, University of Minnesota. Assoc. Mem. ASME.

[3] Professor of Mechanical Engineering, University of Minnesota. Mem. ASME.

[4] Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Heat Transfer Division of The American Society of Mechanical Engineers and presented at the Heat Transfer and Fluid Mechanics Institute, Stanford University, Calif., June 21–23, 1956.

 Manuscript received at ASME Headquarters, May 11, 1956.

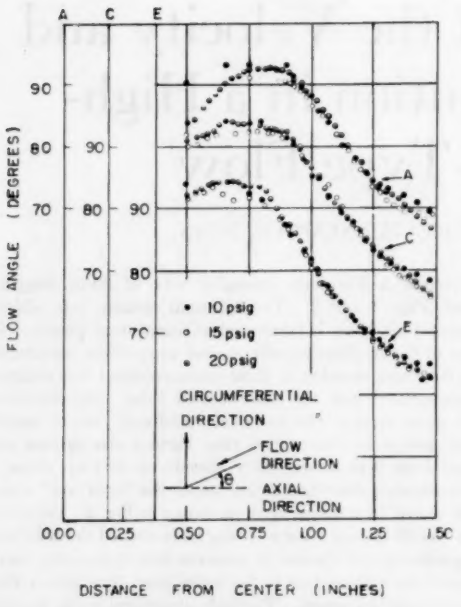


FIG. 3 FLOW ANGLES FOR VORTEX TUBE

The objective of the research program was to obtain a detailed knowledge of the temperature and flow fields in a vortex-type flow. The study was restricted to a vortex which is generated at the nozzle cross section and proceeds in one main direction down the tube to be discharged through the exit at the far end of the tube. To express it another way, the cold-end orifice was completely closed during the present study. After collecting extensive data with the cone-valve-exit geometry, other exit configurations were investigated to determine the influence of this boundary condition on the flow field.

The measurement of the temperature and pressure distribution in a vortex-type flow presents many experimental difficulties. These have been discussed, and the resulting probes for measuring static and total pressures, total temperatures, and flow angles have been described in detail (1, 2). As shown in Fig. 1, opening ports for inserting such measuring probes are located at five axial positions along the tube. The vortex-flow measurements were made at the Rosemount Aeronautical Laboratories of the University of Minnesota where large quantities of clean dry compressed air are available. A calibrated venturi meter arranged upstream of the vortex tube measured the flow rates of the air. After leaving the venturi meter, the air then passed through the vortex tube to be finally discharged into a room with 70 to 80 F temperature.

The experimental program was initiated by flow-visualization studies using wool tufts supported on a fine wire. These studies indicated a negligible radial-velocity component throughout the entire tube. Quantitative measurements were then obtained for an inlet pressure of 10 psig maintaining the same exit-valve opening as shown in Fig. 2. The yaw probe was first inserted into the tube and measurements of the velocity-vector orientation were made at various radial positions and at the axial planes noted in Fig. 3. The calibrated temperature and pressure probes could then be aligned properly in the flow field when making their respective readings.

On completion of the runs with 10 psig inlet pressure, additional data were obtained for inlet pressures of 15 and 20 psig,

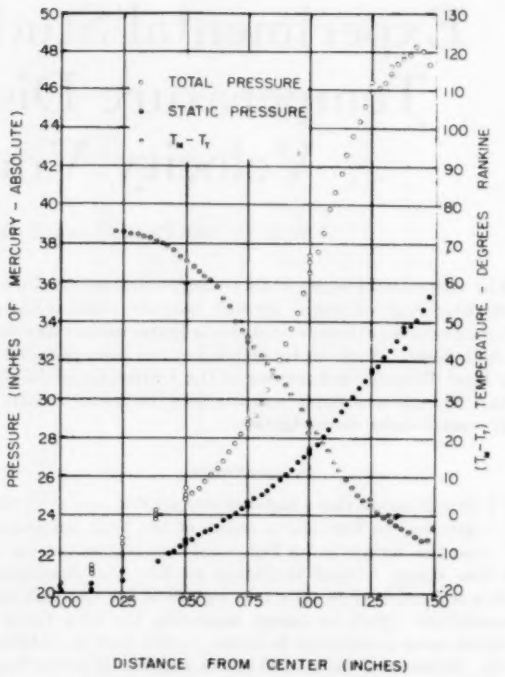


FIG. 4 PRESSURES AND TOTAL TEMPERATURES FOR 20-PSIG INLET PRESSURE (SECTION A)

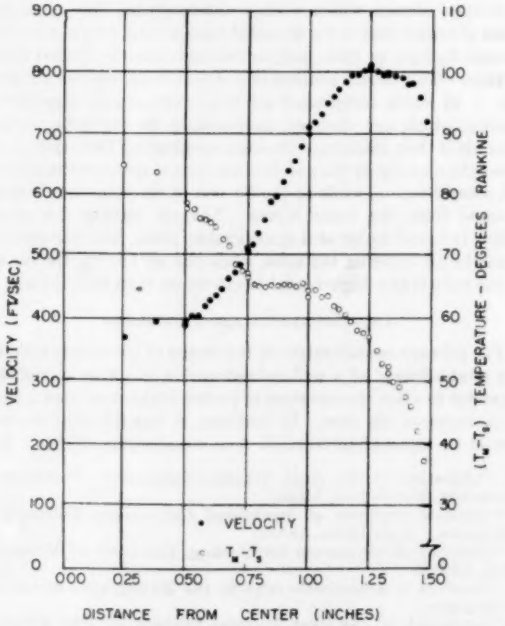


FIG. 5 CALCULATED VELOCITY AND STATIC TEMPERATURES FOR 20-PSIG INLET PRESSURES (SECTION A)

maintaining the same exit opening shown in Fig. 2. To complete the program, measurements were made at 10 psig inlet pressure with two other exit geometries. One geometry was an orifice

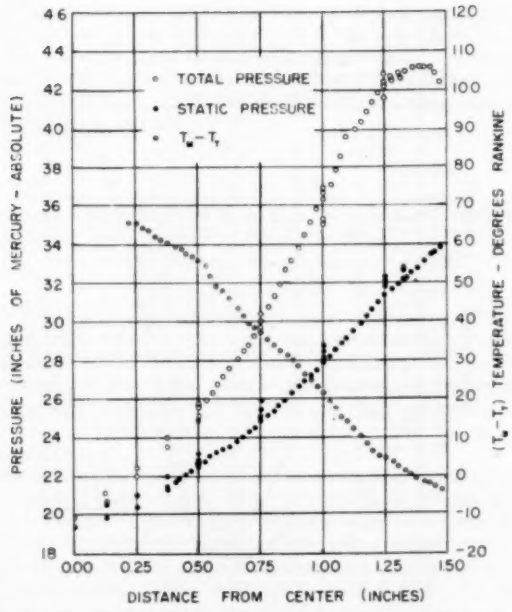


FIG. 6 PRESSURES AND TOTAL TEMPERATURES FOR 20-PSIG INLET PRESSURE (SECTION C)

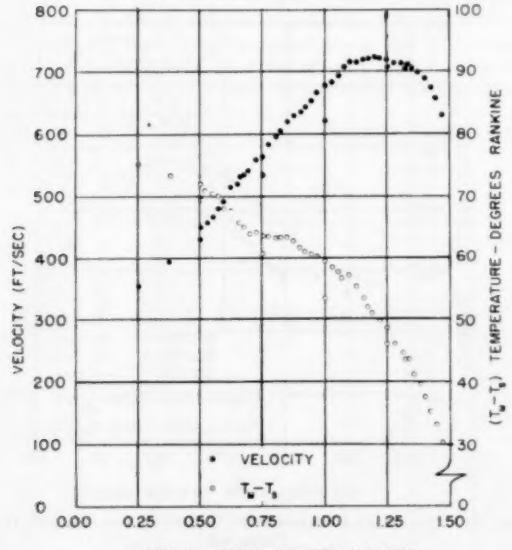


FIG. 7 CALCULATED VELOCITY AND STATIC TEMPERATURES FOR 20-PSIG INLET PRESSURE (SECTION C)

opening of 1.5 in. diam which gives approximately the same area as the cone-shaped valve. The third geometry was a 60-deg nozzle with a 1.5-in-diam exit.

RESULTS

Experimental results were obtained with the geometry shown in Fig. 2 for inlet pressures of 10, 15, and 20 psig. Measured quantities obtained at several cross sections along the tube are

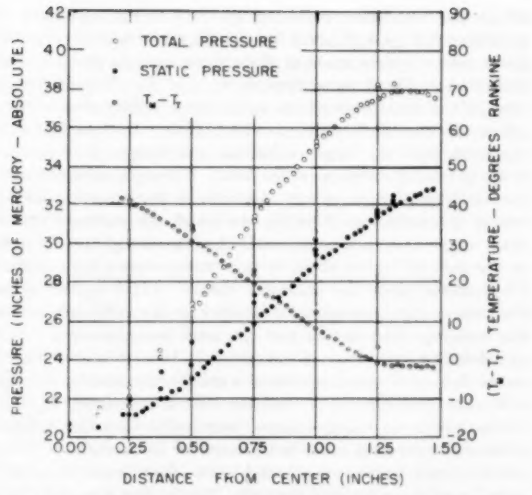


FIG. 8 PRESSURES AND TOTAL TEMPERATURES FOR 20-PSIG INLET PRESSURE (SECTION E)

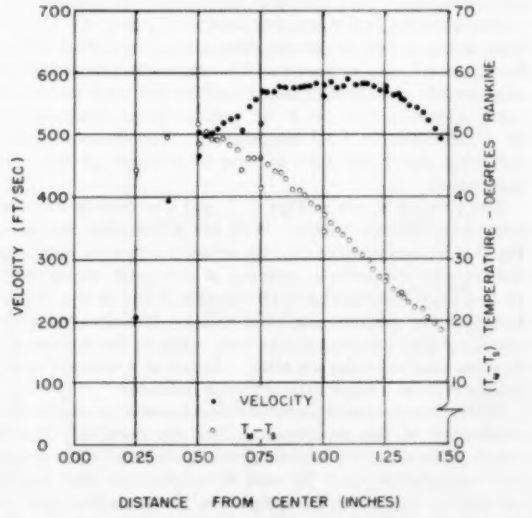


FIG. 9 CALCULATED VELOCITY AND STATIC TEMPERATURES FOR 20-PSIG INLET PRESSURES (SECTION E)

flow angle, total pressures, static pressures, and total temperatures. The flow angle, or more specifically the velocity vector, is shown in Fig. 3 and is found to be quite insensitive to changes in inlet pressure over the range tested. The figure shows that the flow moves in an axial direction toward the discharge value (flow angle less than 90 deg) in the outer regions of the flow near the tube wall. In the flow region in the vicinity of 0.75 in. radius the measurements indicate a flow in the reverse direction, moving from the cone-valve end of the tube toward the nozzle cross section. Inside a radius of 0.5 in. no data are shown since the indication of the probe in this region where it is essentially positioned in its own wake is very erratic. As will be demonstrated, however, the region of principal interest lies outside this inner core.

The local measurements of total pressure P_t, static pressure P_s, and total temperature T_t, obtained with the appropriate probes

set at the orientation indicated by the yaw measurements, are presented in Figs. 4, 6, and 8 for an inlet pressure of 20 psig. The measurements were made at three cross sections along the tube identified in Fig. 1 as sections A, C, and E. The major portion of the measurements is again outside the radius of 0.5 in. where the data are considered to be reliable. In these figures it is apparent that the larger velocities and temperature gradients occur in the outer regions of the flow. The separation of the flow into regions of high energy (high total temperature) and low energy is obvious in these figures which demonstrate that the total temperature is a minimum in the central region. At section A the central region shows a total temperature approximately 75 F lower than the entrance value. (The figures present the temperature depression expressed as the difference between the manifold temperature and the local temperature $T_m - T_t$, and this is a maximum at the center.) It is of interest that the entire flow, with the exception of a small outer annular ring, is at a total temperature lower than the inlet temperature, T_m. It is obvious from an energy-balance standpoint that the deficiency of total temperature must be balanced by an excess of total temperature and this excess is found to be concentrated in a narrow annulus adjacent to the tube wall. As the flow moves down the tube from section A to section E, it is noteworthy that the "separation" effect decreases from a deficiency of 75 F at section A to 65 F at section C, and finally to 40 F at the test station E.

The velocity and static temperature T_s may be calculated from the measured results and these are shown in Figs. 5, 7, and 9. The static-temperature variation across the tube is seen to be considerably smaller than that found for the total temperature, being approximately 50 F at section A as compared with 80 F difference in total temperature. The static-temperature difference across the tube is seen to decrease as the exit is approached.

The velocity shown in Figs. 5, 7, and 9 represents the magnitude of the velocity vector. With the information contained in Fig. 3 it is possible to resolve the velocity into axial and tangential-velocity components, arriving at the result shown in Figs. 10 and 11. The significant information found in Fig. 10 is the fact that the predominant axial velocity is concentrated in an annular region adjacent to the wall, while in the interior of the flow the axial velocities are small. In fact at a radius of approximately 0.75 in. a small reverse flow is indicated.

With the axial and tangential velocities several checks on the consistency of the experimental data are possible. The local values of the density ρ as determined by the local static pressures and temperatures may be used in conjunction with the axial velocity to determine by integration the mass-flow rate down the tube and these values may be compared with the values measured at the venturi meter (Table 1).

TABLE 1 MASS-FLOW BALANCE

Section	Venturi, lbm/sec	Integrated lbm/sec	Error, per cent
	10 Psig Inlet Pressure		
A.....	0.294	0.296	0.7
C.....	0.294	0.293	−0.3
	15 Psig Inlet Pressure		
A.....	0.369	0.390	5.7
B.....	0.366	0.355	−3.0
C.....	0.366	0.386	5.5
E.....	0.381	0.384	0.8
	20 Psig Inlet Pressure		
A.....	0.471	0.498	5.7
C.....	0.471	0.459	−2.5
E.....	0.503	0.459	−8.7

The integrated profiles generally agree within 5 per cent with the measured values with a maximum departure of 8.7 per cent.

Similarly an energy-flow balance may be performed in which the product of the venturi measured flow rate, the specific heat,

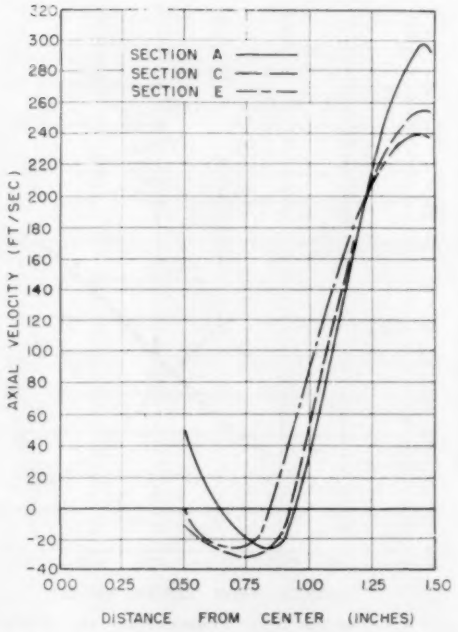


FIG. 10 CALCULATED AXIAL VELOCITY FOR 20-PSIG INLET PRESSURE

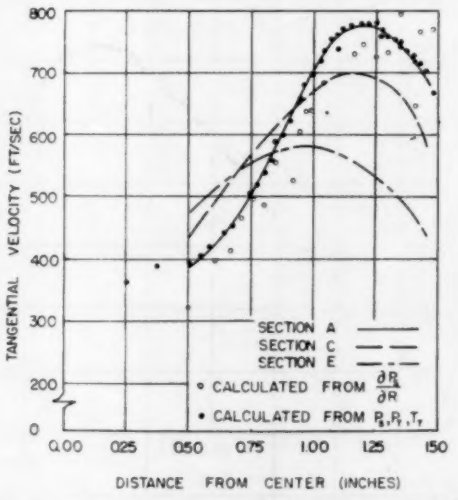


FIG. 11 CALCULATED TANGENTIAL VELOCITY FOR 20-PSIG INLET PRESSURE

TABLE 2 ENERGY BALANCE

Section	Venturi, Btu/sec	Integrated, Btu/sec	Error, per cent
	10 Psig Inlet Pressure		
A.....	36.5	37.3	2.2
C.....	36.5	35.2	−3.6
	15 Psig Inlet Pressure		
A.....	47.5	49.5	4.2
B.....	48.3	45.1	−6.6
C.....	48.3	51.1	5.8
E.....	50.1	50.6	1.0
	20 Psig Inlet Pressure		
A.....	60.7	65.1	7.1
C.....	60.1	57.9	−3.6
E.....	64.5	59.4	−7.8

and the manifold temperature T_m is compared with the integrated energy flow at the several cross sections of the tube (Table 2).

The energy balances agree within 7.8 per cent. One further check on the results can be obtained from a force balance on a fluid particle which relates the static pressure to the tangential velocity by the expression

$$\frac{\partial P_s}{\partial r} = \rho \frac{v^2}{r}$$

The tangential velocity v may thus be calculated from the static-pressure distribution by the relationship

$$v = \sqrt{[(r/\rho)(\partial P_s/\partial r)]}$$

The results of such a calculation are shown in Fig. 11 for 20-psig inlet pressure, where it is apparent that such a calculation agrees with the tangential velocity calculated from the pressures, temperatures, and flow angle. Similar calculations have been carried out for other pressures and location and good agreement also was found. In the interest of space economy they will not be shown here.

An interesting result is obtained by comparing the maximum difference in total temperature encountered at 20-psig inlet pressure with the difference in total temperature encountered in a boundary layer on a flat plate in a flow with a velocity equal to the maximum velocity within the vortex tube. At section A the largest difference in total temperature is 80 F. In a laminar boundary layer in an air stream of 800 fps and 80 F temperature over a flat plate the largest total temperature difference is 11 deg F, and in a turbulent boundary layer 7 F. The energy separation in a boundary layer is apparently much smaller than in vortex flow. On the other hand, experiments (3, 5) on a cylinder in cross flow indicated wall temperatures in the downstream portion which were lower than the static temperature in the oncoming stream. In this case, the separation effect is of the same order as in the vortex tube.

The exit geometry was found to have considerable effect on the flow conditions within the tube and some evidence of this will be discussed shortly. In light of this it is felt that more detailed information on the conditions along the cone-valve surface would be of value. Accordingly, several pressure taps were installed along the cone surface yielding the results shown in Fig. 12 for the 20-psig inlet pressure. The pressure rises along the surface of the cone valve, starting from the low pressures at the center line and ultimately attaining the outside ambient pressure. It is expected from this figure that backward flow occurs within the boundary layer on the cone surface due to the existence of a pressure that increases along the surface.

The effect of varying the exit conditions on the flow distribution may be seen in Fig. 13 where pressure and temperature distribution for the 60-deg nozzle and the 60-deg cone valve are compared. Although the same area opening is used, the flow rate decreased from a value of 0.294 lb_M/sec with the cone valve to 0.250 lb_M/sec with the 60 deg nozzle with the same inlet pressure. This is obviously due to the contraction at the nozzle exit. The measured pressures are higher near the wall and throughout the entire cross section for the nozzle exit. The total temperature is seen to reach a much lower value in the central region for the cone valve. Although not shown, the axial-velocity distribution is quite similar for the two exit geometries, but a marked difference in the tangential velocities is apparent.

Limited data were obtained with a simple orifice opening at the exit (1). The static and total pressure distribution somewhat resemble those described for the 60-deg nozzle, but the total temperature profile is quite different. It is therefore apparent that the influence of the exit boundary condition is felt throughout the entire flow field.

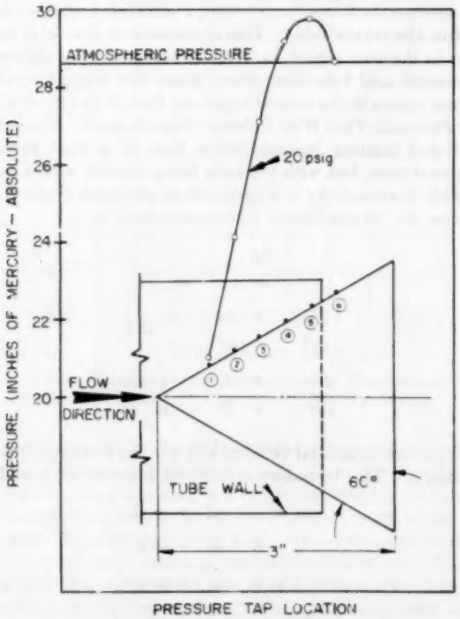


FIG. 12 PRESSURE DISTRIBUTION ALONG SURFACE OF EXIT-CONE VALVE FOR 20 PSIG INLET PRESSURE

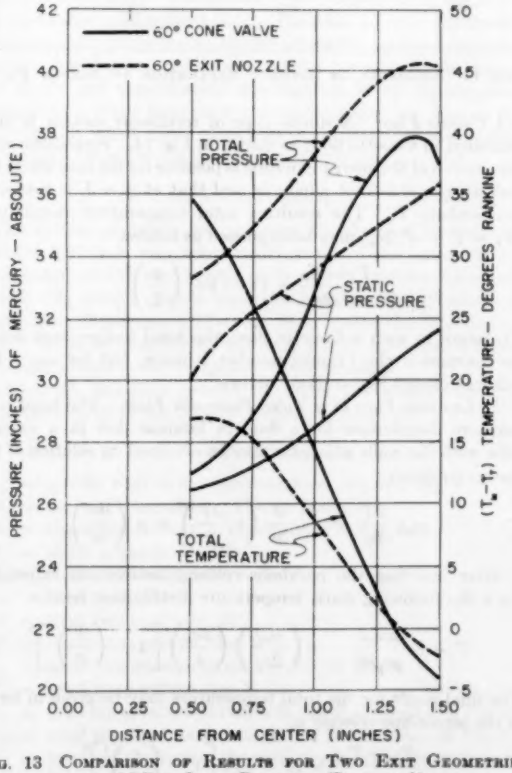


FIG. 13 COMPARISON OF RESULTS FOR TWO EXIT GEOMETRIES FOR 10 PSIG INLET PRESSURE (SECTION A)

Analytical Consideration of Energy Separation

In the steady flow of a fluid without viscosity and heat conductivity there is obviously no possibility for an energy transfer from one stream tube to another since the pressure forces can deliver no work. Consequently, the total energy, or in gases the total temperature, remains constant when initially constant. The situation is different in fluids with viscosity and conductivity even when there is no energy flow through walls. For example, in the boundary layer on an insulated flat plate in a high-velocity gas flow the gas near the wall will have a lower total temperature (lower total energy) than the oncoming free stream, while the flow out away from the wall will be at a higher total temperature than the free stream. Many other rectilinear motions exhibit such an energy separation. It is felt that a study of this energy separation in such flows may well throw some light on the results observed in the vortex tube. For this reason a few examples will be presented herein, assuming incompressible flow:

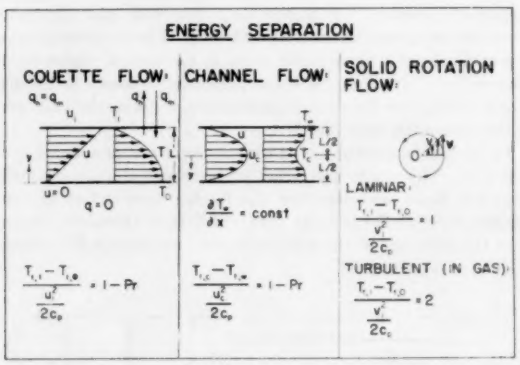


Fig. 14 Examples of Energy Separation in Simple Flows

1 *Couette Flow.* A simple type of rectilinear motion is that exhibited in Couette flow, as shown in Fig. 14. Straightforward integration of the energy equation is possible for the case where the wall at $y = 0$ is kept adiabatic and that at $y = L$ is held at a temperature T_1. The resulting total temperature distribution ($T_t = T + u^2/2c_p$) may be expressed as follows

$$\frac{T_t - T_{t0}}{u_1^2/2c_p} = (1 - \mathrm{Pr})\left(\frac{y}{L}\right)^2 \quad \ldots\ldots [1]$$

Therefore in such a Couette flow, the total temperature would be constant if the Prandtl number is unity, but for any other value an energy "separation" occurs.

2 *Laminar Flow in a Tube, Poiseuille Flow.* The total temperature distribution for a fluid in laminar flow in a circular tube with the walls adiabatic may be obtained on solution of the energy equation

$$\rho c_p u \frac{\partial T}{\partial x} = \frac{k}{r}\frac{\partial}{\partial r}\left(r\frac{\partial T}{\partial r}\right) + \mu\left(\frac{du}{dr}\right)^2 \quad \ldots\ldots [2]$$

After inserting the parabolic-velocity-distribution expression for u the following static temperature distribution results

$$T = 8\frac{\mu u_m x}{\rho c_p R^2} - 4\left(\frac{u_m^2}{2c_p}\right)\left(\frac{\mu c_p}{k}\right)\left[1 - \left(\frac{r}{R}\right)^2\right]^2 \quad \ldots [3]$$

The final result for the total temperature may be given in terms of the center-line velocity u_c

$$\frac{T_t - T_{tw}}{u_c^2/2c_p} = [1 - \mathrm{Pr}]\left[1 - \left(\frac{r}{R}\right)^2\right]^2 \quad \ldots\ldots [4]$$

We again find that for Pr = 1 the total temperature is constant across the tube. The same expression results for the case of laminar flow between parallel walls.

3 *Solid-Body Rotation.* A simple example of rotating flow is laminar solid-body rotation. The velocity distribution varies in this case linearly with the radius

$$v = v_1\left(\frac{r}{R}\right) \quad \ldots\ldots [5]$$

Since there are no shearing stresses present in the assumed flow, the energy equation becomes

$$\frac{\partial}{\partial r}\left(r\frac{\partial T}{\partial r}\right) = 0 \quad \ldots\ldots [6]$$

Imposing the following boundary conditions

$$\left.\begin{aligned} T &= T_w \quad \text{at} \quad r = R \\ \frac{\partial T}{\partial r} &= 0 \quad \text{at} \quad r = 0 \end{aligned}\right\} \quad \ldots\ldots [7]$$

we obtain the solution that $T = T_w$ everywhere in the flow. Therefore there is no heat flow across the wall surface. The resulting dimensionless total temperature difference when referred to the wall becomes

$$\frac{T_{t1} - T_t}{v_1^2/2c_p} = 1 - \left(\frac{r}{R}\right)^2 \quad \ldots\ldots [8]$$

Here we find an important difference from the cases of rectilinear flows, namely, that the "separation" effect is independent of the Prandtl modulus, i.e., we get a variation in total temperature even for a Prandtl number of unity. Thus, for gases with a Prandtl number close to unity, such a "separation" is expected to be small for rectilinear flows; but if a rotation is superimposed, it would appear that considerable variation in total temperature could be established. It is obvious that the central total temperatures are the lowest in the tube, a situation analogous to that found in the vortex tube. This occurrence of low total temperatures in the core region for the solid-rotation case differs from the channel and tube-flow cases, where the highest total temperature occurs in the central region for flow of air (Pr = 0.72).

4 *Poiseuille Flow With Rotation Superimposed.* Consider the established laminar incompressible flow of a fluid through a cylindrical tube, but with the tube being rotated with a velocity v_1. This is essentially a superposition of Cases 2 and 3. For this case the Navier-Stokes equations reduce to

$$\left.\begin{aligned} \frac{\partial p}{\partial r} &= \rho\frac{v^2}{r} \\ \frac{d^2v}{dr^2} + \frac{1}{r}\frac{dv}{dr} - \frac{v}{r^2} &= 0 \\ \frac{d^2u}{dr^2} + \frac{1}{r}\frac{du}{dr} &= \frac{1}{\mu}\frac{\partial p}{\partial x} \end{aligned}\right\} \quad \ldots\ldots [9]$$

where v is the tangential velocity and u is the axial velocity in the direction x. The boundary conditions imposed on v and u are

$$\left.\begin{aligned} v &= v_1 \quad \text{at} \quad r = R \\ v &= 0 \quad \text{at} \quad r = 0 \\ u &= 0 \quad \text{at} \quad r = R \\ \frac{du}{dr} &= 0 \quad \text{at} \quad r = 0 \end{aligned}\right\} \quad \ldots\ldots [10]$$

The velocity components as obtained by solution of these equations are

$$\left.\begin{aligned} u &= 2u_m[1 - (r/R)^2] \\ v &= v_1(r/R) \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots [11]$$

The energy equation for this case is

$$\rho c_p u \frac{\partial T}{\partial x} = \frac{k}{r}\frac{\partial}{\partial r}\left(r\frac{\partial T}{\partial r}\right) + \mu\left(\frac{du}{dr}\right)^2 \quad \ldots\ldots [12]$$

The following boundary conditions may be imposed

$$\left.\begin{aligned} \frac{\partial T}{\partial r} &= 0 \quad \text{at} \quad r = 0 \\ \frac{\partial T}{\partial r} &= 0 \quad \text{at} \quad r = R \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots [13]$$

The solution is precisely that given for pipe flow, Case 2

$$T = \frac{8\mu u_m}{\rho c_p R^2}x - 4\left(\frac{u_m^2}{2C_p}\right)\left(\frac{\mu C_p}{k}\right)\left[1 - \left(\frac{r}{R}\right)^2\right]^2 \quad \ldots [14]$$

If the static temperature T is converted to the total temperature and is made dimensionless, using the center-line axial velocity u_c, the following equation results

$$\frac{T_t - T_{tw}}{u_c^2/2c_p} = [1 - \mathrm{Pr}]\left[1 - \left(\frac{r}{R}\right)^2\right]^2 - \left(\frac{v_1}{u_c}\right)^2\left[1 - \left(\frac{r}{R}\right)^2\right] \quad \ldots\ldots [15]$$

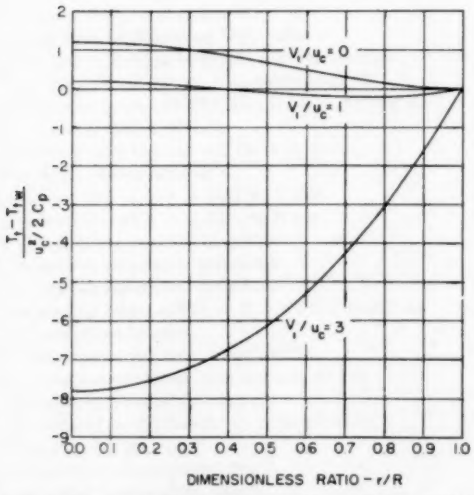


FIG. 15 TOTAL TEMPERATURE DISTRIBUTION POISEUILLE FLOW THROUGH A ROTATING CIRCULAR TUBE (PR = 0.7)

The result is shown in Fig. 15 for several values of the parameter (v_1/u_c) and for a Prandtl number of 0.7. Here it is plainly indicated that the superimposing of a rotational velocity on the normal Poiseuille flow creates lower total temperatures in the core of the flow, the effect becoming larger with increasing rotation. Since extremely large ratios of circumferential to axial velocity occur within the vortex tube, one would predict very low total temperatures in the core regions. The axial component of velocity prevents this separation from ever reaching the value obtainable with pure rotation.

5 *Turbulent Flow, Vortex Motion.* It is hypothesized (4, 6, 7) that the equilibrium temperature distribution in a highly turbulent gas flow in which pressure gradients exist normal to the flow direction should correspond to an isentropic variation with pressure; that is, if a small mass of fluid at temperature T_1 and pressure P_{s1} moves to a new position where the pressure is P_{s2} the small mass will take up a temperature T_2 which is given by

$$T_2 = T_1\left(\frac{P_{s2}}{P_{s1}}\right)^{\frac{k-1}{k}}$$

If the temperature of the fluid already at the new position corresponds to this temperature T_2, no heat transfer will occur; but if it is different, then a heat transfer will result and tend to establish the temperature T_2. This is essentially the condition existing in the atmosphere with strong turbulent mixing.

If we accept this hypothesis and consider a turbulent rotating flow with solid-body rotation satisfying the necessary force balance, we arrive at the following expression for the total temperature difference between the center line and the wall positions

$$\frac{T_{ti} - T_{t0}}{v_1^2/2c_p} = 2 \quad \ldots\ldots\ldots\ldots [16]$$

Even when it is realized that the foregoing model makes some severe assumptions it is felt that it throws light on the vortex-tube study. Equation [16] indicates a greater separation effect in turbulent rotational flow than in laminar rotating flow. The value of the dimensionless parameter on the left-hand side of Equation [16], if v_1^2 is taken to be the maximum velocity in the vortex tube, turns out to be from 1.3 to 1.8 in the experimental investigation, a value higher than predicted for laminar flow, but lower than predicted in turbulent flow. The axial velocity apparently keeps the value below the maximum attainable, as was mentioned earlier. Another point of correspondence between the turbulent-flow model and experimental investigation is the experimental evidence of a static temperature gradient across the flow in the outer regions which more closely approximates that assumed in the turbulent case than in the laminar calculation.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the efforts of Roland and Richard Birkebak who constructed many of the pressure and temperature probes used in this study and in addition have performed the measurements reported herein. The Mechanical Engineering Department Machine Shop under the direction of Robert Robarge constructed the 3-in. vortex tube, and Francis J. Bradac helped in the design and construction.

This work has been carried on under the sponsorship and financial support of the Office of Ordnance Research.

NOMENCLATURE

- c_p = specific heat at constant pressure, Btu/lb$_M$ deg F
- k = thermal conductivity, Btu/hr ft deg F
- L = characteristic length, ft, identified in Fig. 14
- P_s = static pressure, lb$_F$/ft^2
- P_t = total pressure, lb$_F$/ft^2
- r = local radius, ft
- R = total radius of tube, ft
- T = static temperature, deg R
- T_m = total temperature in inlet manifold, deg R
- T_t = total temperature, deg R
- T_{ti} = total temperature at radius R, deg R, identified in Fig. 14
- T_{t0} = total temperature at center, deg R, identified in Fig. 14
- T_{tw} = total temperature at the wall, deg R, identified in Fig. 14
- u = local velocity in the x direction, ft/sec

u_m = mean velocity in the x direction, ft/sec
u_e = velocity in the x direction calculated at the centerline position, ft/sec
v = velocity in the tangential direction, ft/sec
v_i = tangential velocity at the radius R, ft/sec, identified in Fig. 14
x = co-ordinate in main flow direction, ft
y = co-ordinate normal to main flow direction, ft
ρ = mass density, lb_M/ft^3
μ = viscosity, $lb_M/hr\ ft$
Pr = Prandtl number = $\frac{\mu c_p}{k}$

BIBLIOGRAPHY

1 "Investigation of the Energy Distribution in a High Velocity Vortex Type Flow," by E. R. G. Eckert and J. P. Hartnett, University of Minnesota Heat Transfer Laboratory Technical Report No. 3, June, 1955.

2 "Experimental Study of the Velocity and Temperature Distribution in a High Velocity Vortex Type Flow,'" by E. R. G. Eckert and J. P. Hartnett, University of Minnesota Heat Transfer Laboratory Technical Report No. 6, September, 1955.

3 "Messungen der Temperatur-Verteilung auf der Oberflaeche schnell angestroemter Koerper," by E. R. G. Eckert and W. Weise, *Forschung auf dem Gebiete des Ingenieurwesens*, vol. 13, 1942, pp. 248–254.

4 "Friction Laws and Energy Transfer in Circular Flow," by R. Kassner and E. Knoernschild, Technical Report No. F-TR2198-ND, Wright Patterson Air Force Base, Ohio, March, 1948.

5 "Experiments on Aerodynamic Cooling," by L. F. Ryan, Mitteilungen Institute Aerodynamik, Eidgen. Tech. Hochschule, Zurich, Switzerland, No. 18, 1951, pp. 7–50.

6 "Die Wirkungensweise des Ranque-Wirbelrohres," by F. Schultz-Grunow, (The Operation of the Ranque Vortex Tube), *Kaeltetechnik*, Band 2, No. 11, 1950, pp. 273–274.

7 "On the Theory of the Ranque-Hilsch Cooling Effect," by J. J. Van Deemter, *Applied Scientific Research* (Netherlands) section A, vol. 3, March, 1952, pp. 174–196.

Solving the Melting Problem Using the Electric Analogy to Heat Conduction

By D. R. OTIS,[1] SAN DIEGO, CALIF.

The one-dimensional, heat-conduction equation with melting at one boundary is transformed to a form suitable for simulation on an analog computer. The transformed equation is identical with the heat-conduction equation with internal heat generation and a nonmelting boundary. This allows simulation of the melting process continuously with time rather than stepwise, and the number of "cells" in the heat-conduction-circuit analog remains constant throughout the solution.

Nomenclature

The following nomenclature is used in the paper:

a = slab thickness, ft
b = length of electric wire, ft
c = specific heat, Btu/lb deg F
C = capacitance per cell, microfarads
C' = capacitance per foot, microfarads/ft
e_n = defined by Equation [37], volts
E = potential in the ideal electrical system, volts
$\bar{E}$ = E/E_{max}, dimensionless
E_L = defined by Equation [32], volts
E_n = potential of nth cell, volts
E_s = defined by Equation [31], volts
I_n = current leakage from nth cell, microamps
I' = current leakage per foot, microamps/ft
k = thermal conductivity, Btu ft/sq ft sec deg F
L = heat of fusion, Btu/lb
n = designates the nth cell ($n = 0, 1, \ldots . N$)
N = $b/\Delta y$, dimensionless
q_a = heat flux at $x = a$, Btu/sq ft sec
q_s = heat flux at $x = s$, Btu/sq ft sec
R = resistance per cell, megohms
R_a = scaling resistance, megohms
R_L = scaling resistance, megohms
R_n = scaling resistance ($n = 0, 1, .. N$), megohms
R_s = scaling resistance
R' = resistance per foot, megohms/ft
s = position of the melting surface, $s(t)$, ft
t = time in thermal system, sec
$\bar{t}$ = defined by Equation [9], dimensionless
T = temperature, deg F
T_m = melting temperature, deg F
T_0 = initial temperature, deg F
$\bar{T}$ = defined by Equation [10], dimensionless
x = length co-ordinate, ft
$\bar{x}$ = defined by Equation [7], dimensionless
y = length co-ordinate, ft
$\bar{y}$ = y/b, dimensionless
α = thermal diffusivity, sq ft/sec
ρ = density, pcf
η = variable of integration, sec
τ = computing time, sec
$\bar{\tau}$ = τ/τ_{max}, dimensionless
Δy = incremental length, ft
$\Delta \bar{y}$ = $\Delta y/b$, dimensionless

The subscript "max" is associated with constant scaling factors which are chosen arbitrarily to suit the physical problem and the computing equipment.

Introduction

The scarcity of solutions reported in the literature lends testimony to the difficulty of the melting problem. Several rather restricted solutions are presented by Carslaw and Jaeger (1).[2] The equations were formulated in a very general form by Landau (2) who gave analytic solutions for several limiting cases and a numerical solution for the case of a semi-infinite slab heated by a constant heat flux. The assumptions required in these solutions often proved to be too restrictive, and the need for a more general method of solution became evident. A circuit was developed at Convair, San Diego, for simulating the melting problem using the analogy between the flow of heat and the flow of electric current. The circuit simulated a slab of finite thickness with variable thermal properties, and melting could occur at one surface. The boundary conditions were completely general being limited only by the equipment available to simulate them. The system used to simulate the variation in thermal properties has been reported elsewhere (3), and this paper is concerned with the method of handling the melting-boundary condition.

There are two distinct methods for simulating heat-conduction problems, and these methods utilize circuits which are sometimes referred to as the "active network" and the "passive network." With the active network, the heat-conduction equation is approximated by a set of ordinary differential equations which are solved using an electronic differential analyzer (4). With the passive network, use is made of the analogy between the flow of electric current and the conduction of heat (5). This paper is concerned with a "passive network."

Usually when one thinks of simulating the melting problem by electric analogy, it is thought that the electric circuit must consist of many "cells" (each cell corresponding to some thickness of the heated material) in which the melting is represented by removing cells from the circuit as the melting proceeds. This would require a system of relays and associated equipment, and would result in a stepwise approximation to the melting process. The method presented in this paper solves the melting problem not by removing cells, but by mathematically shrinking the co-ordinate system as melting occurs during the solution to the problem. The advantages of this method are twofold: (*a*) A minimum number of cells is required since their number will be constant throughout the solution; and (*b*) the melting process is represented continuously with time rather than stepwise.

---

[1] Thermodynamics Group, Convair.

Contributed by the Heat Transfer Division of The American Society of Mechanical Engineers and presented at the Heat Transfer and Fluid Mechanics Institute, Stanford, Calif., June 21–23, 1956.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, May 11, 1956.

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

Analysis

The Heat-Conduction Equation. The heat-conduction equation for one-dimensional heat flow in rectangular co-ordinates with constant thermal properties and no heat source is

$$\frac{\partial T}{\partial t} = \alpha \frac{\partial^2 T}{\partial x^2} \quad [1]$$

The boundary conditions considered in this paper are

$$t = 0 \qquad T = T_0 \qquad s(t) = 0 \quad [2]$$

$$x = s(t) \qquad -k \frac{\partial T}{\partial x} + \rho L \frac{ds}{dt} = q_s \quad [3]$$

$$x = a \qquad -k \frac{\partial T}{\partial x} = q_a \quad [4]$$

The functions q_s and q_a represent the heat fluxes at the boundaries, and no restrictions will be imposed upon these quantities. The position of the heated surface is represented by $s(t)$ and can only be determined with the following specification

$$\frac{ds}{dt} = 0 \quad \text{when} \quad T_{x=s} < T_m \quad [5]$$

$$\frac{ds}{dt} \geq 0 \quad \text{when} \quad T_{x=s} = T_m \quad [6]$$

The temperature of the melting surface cannot exceed T_m. Following a procedure similar to that of Landau (2), the moving boundary can be eliminated by introducing a new dimensionless variable defined by

$$\bar{x} = \frac{x - s}{a - s} \quad [7]$$

Equation [1] becomes

$$\frac{\partial T}{\partial t} = \frac{\alpha}{(a-s)^2} \frac{\partial^2 T}{\partial \bar{x}^2} + \frac{1 - \bar{x}}{a - s} \frac{ds}{dt} \frac{\partial T}{\partial \bar{x}} \quad [8]$$

In order that the variable coefficient be eliminated from the first term on the right-hand side of Equation [8], a new dimensionless time variable is introduced which is defined by

$$\bar{t} = \frac{a^2}{t_{\max}} \int_0^t \frac{d\eta}{(a-s)^2} \quad [9]$$

So that the dependent variable also will be dimensionless, the following equation defines a new temperature variable

$$\bar{T} = \frac{T - T_0}{T_m - T_0} \quad [10]$$

Equation [1] can now be written as follows

$$\frac{\partial \bar{T}}{\partial \bar{t}} = \frac{t_{\max}\alpha}{a^2} \frac{\partial^2 \bar{T}}{\partial \bar{x}^2} + \frac{1 - \bar{x}}{a - s} \frac{ds}{dt} \frac{\partial \bar{T}}{\partial \bar{x}} \quad [11]$$

The boundary conditions are

$$\bar{t} = 0 \qquad \bar{T} = 0 \qquad s = 0 \quad [12]$$

$$\bar{x} = 0 \qquad -k \frac{T_m - T_0}{a - s} \frac{\partial \bar{T}}{\partial \bar{x}} + \rho L \frac{a^2}{t_{\max}(a-s)^2} \frac{ds}{d\bar{t}} = q_s \quad [13]$$

$$\bar{x} = 1 \qquad -k \frac{T_m - T_0}{a - s} \frac{\partial \bar{T}}{\partial \bar{x}} = q_a \quad [14]$$

Now Equation [11] is simply the equation for heat conduction where the last term represents a distributed heat sink (since the temperature gradient is negative the last term corresponds to heat removal). With the moving boundary eliminated from the boundary conditions, the problem is reduced to that of solving the heat-conduction equation for a nonmelting solid with internal heat absorption.

The Electrical-Circuit Equation. Imagine an electric wire with a series resistance of R' megohms per foot, a shunt-current leakage of I' microamperes per foot and a shunt capacitance of C' microfarads per foot. For such a system, an equation can be derived for the potential at any point along the wire, and at any time, by applying Kirchoff's first law to an infinitesimal segment of the wire. This is nothing more than a special case of "the telegrapher's equation" (6) where the series inductance of the wire is zero, and the equation can be written as follows

$$\frac{\partial E}{\partial \tau} = \frac{1}{R'C'} \frac{\partial^2 E}{\partial y^2} + \frac{I'}{C'} \quad [15]$$

The analog circuit is arranged to satisfy the following boundary conditions

$$\tau = 0 \qquad E = 0 \quad [16]$$

$$y = 0 \qquad -\frac{1}{R'} \frac{\partial E}{\partial y} = \frac{E_s - E_{y=0}}{R_s} + \frac{E_L - E_{y=0}}{R_L} \quad [17]$$

$$y = b \qquad -\frac{1}{R'} \frac{\partial E}{\partial y} = \frac{E_a - E_{y=b}}{R_a} \quad [18]$$

By introducing the dimensionless variables

$$\bar{E} = \frac{E}{E_{\max}} \quad [19]$$

$$\bar{\tau} = \frac{\tau}{\tau_{\max}} \quad [20]$$

$$\bar{y} = \frac{y}{b} \quad [21]$$

Equation [15] and the Boundary Conditions [16], [17], and [18] become

$$\frac{\partial \bar{E}}{\partial \bar{\tau}} = \frac{\tau_{\max}}{b^2 R'C'} \frac{\partial^2 \bar{E}}{\partial \bar{y}^2} + \frac{\tau_{\max}}{E_{\max}} \frac{I'}{C'} \quad [22]$$

$$\bar{\tau} = 0 \qquad \bar{E} = 0 \quad [23]$$

$$\bar{y} = 0 \qquad -\frac{1}{bR'} \frac{\partial \bar{E}}{\partial \bar{y}} = \frac{\bar{E}_s - \bar{E}_{\bar{y}=0}}{R_s} + \frac{\bar{E}_L - \bar{E}_{\bar{y}=0}}{R_L} \quad [24]$$

$$\bar{y} = 1 \qquad -\frac{1}{bR'} \frac{\partial \bar{E}}{\partial \bar{y}} = \frac{\bar{E}_a - \bar{E}_{\bar{y}=1}}{R_a} \quad [25]$$

Conditions of Similarity. It is easily seen that Equations [11] and [22] are of identical form, and so are their corresponding boundary conditions. The two equations can be related by

$$\bar{T} = \bar{E} \quad [26]$$

$$\bar{t} = \bar{\tau} \quad [27]$$

$$\bar{x} = \bar{y} \quad [28]$$

$$\frac{t_{\max}\alpha}{a^2} = \frac{\tau_{\max}}{b^2 R'C'} \quad [29]$$

$$\frac{1 - \bar{x}}{a - s} \frac{ds}{d\bar{t}} \frac{\partial \bar{E}}{\partial \bar{y}} = \frac{\tau_{\max}}{E_{\max}} \frac{I'}{C'} \quad [30]$$

The foregoing conditions require that the boundary conditions be related as follows

$$\frac{a - s}{k(T_m - T_0)} q_s = bR' \left[\frac{\bar{E}_s - \bar{E}_{\bar{y}=0}}{R_s} \right] \quad [31]$$

$$-\frac{a-s}{k(T_m-T_0)}\left[\frac{\rho La^2}{t_{max}(a-s)^2}\frac{ds}{d\bar{t}}\right]=bR'\left[\frac{\bar{E}_L-\bar{E}_{\bar{y}=0}}{R_L}\right] \quad [32]$$

$$\frac{a-s}{k(T_m-T_0)}q_s=bR'\left[\frac{\bar{E}_s-\bar{E}_{\bar{y}=1}}{R_s}\right] \quad [33]$$

Using Equations [26] through [33], a solution to Equation [22] can be interpreted as a solution to Equation [11]. Therefore, the problem is now resolved to one of finding a means for solving Equation [22].

The Analog Circuit

The equation for the idealized electrical system can be instrumented approximately using the so-called "lumped constant" circuit which is shown in Fig. 1. The wire is divided into sections of length Δy; the resistance, capacitance, and shunt-current leakage for these sections are "lumped" such that

$$R=R'\Delta y=R'\left(\frac{b}{N}\right) \quad [34]$$

$$C=C'\Delta y=C'\left(\frac{b}{N}\right) \quad [35]$$

$$I=I'\Delta y=I'\left(\frac{b}{N}\right) \quad [36]$$

The circuitry associated with each section of wire is called a "cell." The number of cells in the circuit is given by $N+1$.

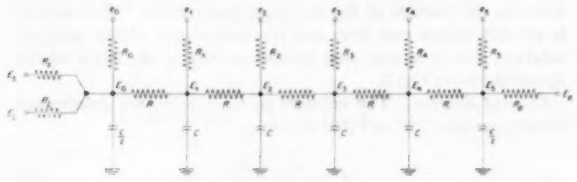


Fig. 1

Now the question arises as to how one can supply all the necessary voltages and choose the circuit constants for this lumped constant circuit, and this will be discussed in the following sections.

Definition of R and C. R and C are determined by combining Equations [29], [34], and [35] obtaining

$$RC=\frac{a^2\tau_{max}}{N^2\alpha t_{max}} \quad [37]$$

The choice of R and C is somewhat arbitrary, since it is only necessary that the product satisfy Equation [37].

The Boundary Conditions. The condition for zero time is easily satisfied by requiring that all capacitors be discharged initially. This requires some sort of switching system to discharge the capacitors after each run.

The voltages E_s and E_a (which correspond to the heat fluxes) are obtained from Equations [31] and [33] with the aid of Equation [34]

$$\bar{E}_s=\bar{E}_{\bar{y}=0}+\frac{R_s(a-s)}{NRk(T_m-T_0)}q_s \quad [38]$$

$$\bar{E}_a=\bar{E}_{\bar{y}=1}+\frac{R_a(a-s)}{NRk(T_m-T_0)}q_a \quad [39]$$

Let it be required that the following voltages will be supplied for E_s and E_a

$$E_s=E_{\bar{y}=0}+\left(\frac{a-s}{a}\right)\frac{q_s}{q_{smax}}E_{max} \quad [40]$$

$$E_a=E_{\bar{y}=1}+\left(\frac{a-s}{a}\right)\frac{q_a}{q_{amax}}E_{max} \quad [41]$$

where the subscript "max" refers to arbitrarily chosen constants. These two equations then define R_s and R_a for, by substitution of Equation [40] into [38] and Equation [41] into [39], one obtains

$$R_s=\frac{NRk(T_m-T_0)}{aq_{smax}} \quad [42]$$

$$R_a=\frac{NRk(T_m-T_0)}{aq_{amax}} \quad [43]$$

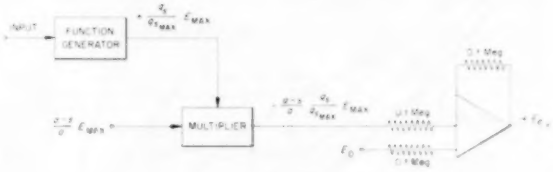


Fig. 2

The instrumentation of Equation [40] is shown in Fig. 2.[3] The function generator generates the arbitrary heat-flux function, and the input to the function generator can be driven by a voltage which is a function of time, or surface temperature, or both. The output of the function generator is then multiplied by the factor $(a-s)/a$ (which is yet to be discussed) and added to $E_{\bar{y}=0}$ giving the desired voltage for application to the lumped constant circuit. The same procedure is used in obtaining E_a.

Relating Real Time and Machine Time. Real time is related to the transformed time by Equation [9] which can be rewritten as

$$t=t_{max}\int_0^{\bar{t}}\left(\frac{a-s}{a}\right)^2 d\eta \quad [44]$$

The transformed time is related to machine time by Equations [20] and [27], and this relationship is

$$\tau=\tau_{max}\bar{t} \quad [45]$$

Then Equation [44] can be written

$$t=\frac{t_{max}}{\tau_{max}}\int_0^{\tau}\left(\frac{a-s}{a}\right)^2 d\eta \quad [46]$$

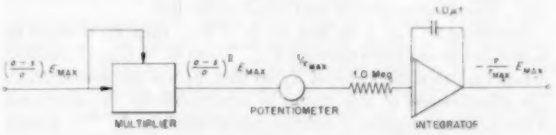


Fig. 3

In order to obtain real time from machine time an integration is required, and the circuit is shown in Fig. 3. The output of this circuit is used to drive the X-co-ordinate of the plotting board so that the results can be plotted directly as a function of real time. It also is used to drive function generators for time-dependent boundary conditions.

Definition of E_L. The voltage E_L is related to the rate of melting and to the position of the melting surface. These are both

[3] Analog computing techniques are described in reference (7).

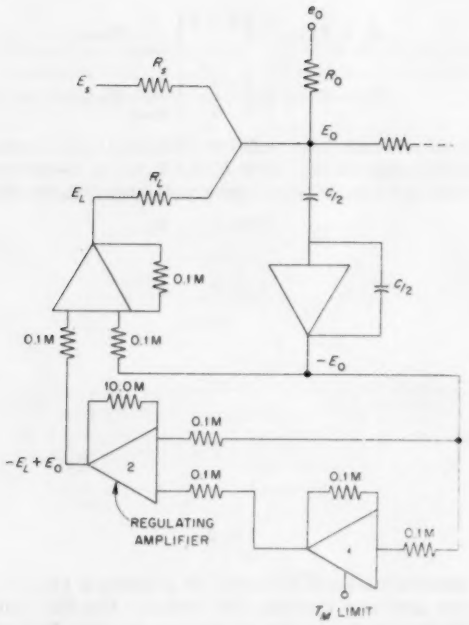


FIG. 4

unknowns. In fact, it is the rate of melting and the position of the melting surface that are desired as the solution to the problem. E_L is fixed by the requirement that the temperature can never exceed the melting temperature. A voltage limiter is connected to the first cell of the circuit to limit the maximum voltage to a value corresponding to the melting temperature ($\bar{E} = 1$). This circuit is shown in Fig. 4.

At this point, it would be well to point out a difference between Figs. 1 and 4. In Fig. 1, one terminal of each capacitor is connected to ground. In Fig. 4, one end of the capacitor for the first cell is connected to the grid of a high gain d-c amplifier. The feedback on the amplifier is a capacitor equal in size to the one in the circuit shown in Fig. 1. This simply provides a means of obtaining the voltages at each cell without influencing the behavior of the cell. The grid of the amplifier is effectively at ground potential (within a few millivolts), so for all practical purposes the lower end of the capacitor is grounded. The output of the amplifier then reads the negative of the voltage of the first cell. This same system is used for all of the cells.

Returning to Fig. 4, so long as $E_0 < E_{max}$ then E_L is equal to E_0 and no current flows through R_L. When $E_0 = E_{max}$ the limiter circuit regulates the value of E_L so that E_0 will not exceed E_{max}. The effectiveness of this circuit is determined by the sharpness of the limit on amplifier 1 and the gain of amplifier 2. The higher the gain of the regulating amplifier the better is the regulation, but it was found that for gains greater than 100 the circuit tended to become unstable.

The output of the regulating amplifier is quite important since it is a measure of the rate of melting. Also, the integral of the rate of melting with respect to time gives the position of the melting surface. From Equations [19], [32], and [34] the output of the regulating amplifier is

$$-E_L + E_{y=0} = -E_L + E_0 = + \frac{R_L a \rho L}{NRk(T_m - T_0)t_{max}} \frac{a}{a-s} \frac{ds}{dl} E_{max} \quad [47]$$

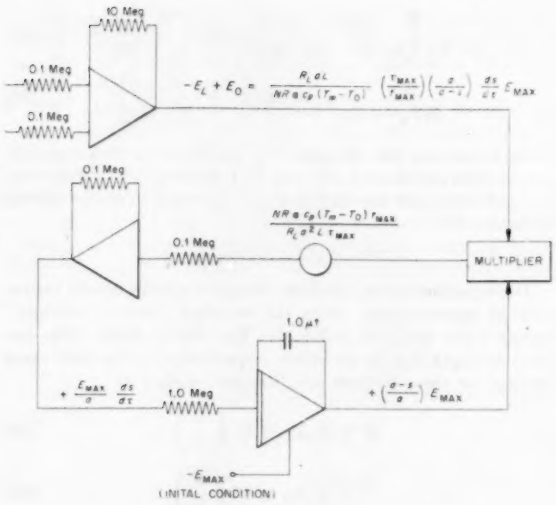


FIG. 5

$$= + \frac{R_L a L}{NRac_p(T_m - T_0)} \frac{\tau_{max}}{t_{max}} \frac{a}{a-s} \frac{ds}{d\tau} E_{max} \quad [48]$$

The circuit shown in Fig. 5 solves this equation for $(a - s)/a$ which is the fraction of the slab that is unmelted. This output is plotted versus real time and represents part of the problem solution, and it is also used simultaneously as an input to the circuit shown in Fig. 2.

Current Sources. The voltages $e_0, e_1, \ldots . e_N$ are determined from Equations [30] and [36] where

$$I_n = \frac{e_n - E_n}{R_n} \quad [49]$$

Then

$$\frac{e_n - E_n}{R_n} = C \frac{E_{max}}{\tau_{max}} \frac{1-(n/N)}{a-s} \frac{ds}{dl} \frac{\partial \bar{E}}{\partial \bar{y}} \quad [50]$$

$$= \frac{a^2 E_{max}}{RN^2 \alpha l_{max}} \frac{1-(n/N)}{a-s} \frac{ds}{dl} \frac{\partial \bar{E}}{\partial \bar{y}} \quad [51]$$

First of all, it is necessary to make some approximation for the potential gradient $(\partial \bar{E}/\partial \bar{y})$. The following approximations were used:

For $n = 1, 2, \ldots . N$

$$\left(\frac{\partial \bar{E}}{\partial \bar{y}}\right)_n \approx \frac{\bar{E}_{n+1} - \bar{E}_{n-1}}{2\Delta \bar{y}} = \frac{N}{2}(\bar{E}_{n+1} - \bar{E}_{n-1}) \quad [52]$$

For $n = 0$

$$\left(\frac{\partial \bar{E}}{\partial \bar{y}}\right)_0 \approx \frac{\bar{E}_1 - \bar{E}_0}{\Delta \bar{y}} = N(\bar{E}_1 - \bar{E}_0) \quad [53]$$

It would be difficult to defend the approximation for $n = 0$, and no attempt will be made to do so. It was always intended that this would be improved; however, the pressing nature of the problem and the agreement with the check solutions discouraged further refinement.

Using the foregoing approximations for the gradients, the current-source voltages are defined by:

For $n = 0$

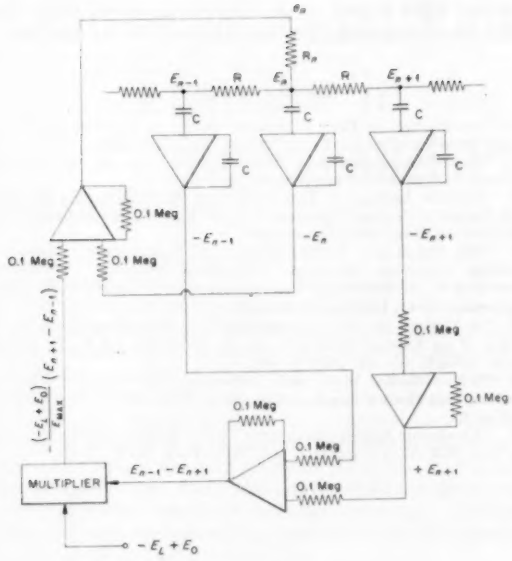


FIG. 6

$$e_0 = E_0 + \frac{R_0 a}{RN\alpha t_{max}} \left(\frac{a}{a-s}\right) \frac{ds}{dl} (E_1 - E_0) \quad [54]$$

For $n = 1, 2, \ldots N-1$

$$e_n = E_n + \frac{R_n a[1-(n/N)]}{2RN\alpha t_{max}} \left(\frac{a}{a-s}\right) \frac{ds}{dl} (E_{n+1} - E_{n-1}) \quad [55]$$

For $n = N$

$$e_N = E_N \quad [56]$$

Equation [56] means that no current source is required on the last cell.

Now let it be required that the following voltages will be supplied for e_n:

For $n = 0$

$$e_0 = E_0 + \frac{(-E_L + E_0)(E_1 - E_0)}{E_{max}} \quad [57]$$

For $n = 1, 2, \ldots N-1$

$$e_n = E_n + \frac{(-E_L + E_0)(E_{n+1} - E_{n-1})}{E_{max}} \quad [58]$$

In order for Equations [57] and [58] to be consistent with Equations [54] and [55], R_n must be defined as follows:

For $n = 0$

$$R_0 = \frac{R_L L}{c_p(T_m - T_0)} \quad [59]$$

For $n = 1, 2, \ldots N-1$

$$R_n = \frac{2R_L L}{[1-(n/N)]c_p(T_m - T_0)} \quad [60]$$

The circuit for Equation [58] is shown in Fig. 6.

RESULTS

The circuit described in the preceding section was used in conjunction with a wide variety of boundary conditions; however, the only comparison that was made with a numerical solution was for the case of a constant heat flux applied at one surface. The other surface was insulated. This was compared with the work of Landau for a constant heat flux at the surface of a semi-infinite slab. The two cases are not exactly the same, and good agreement could only be expected in the early stages of melting while the insulated surface of the analog solution is not affected appreciably. That is, as long as the case simulated on the computer is acting like a semi-infinite slab, it should agree with Landau's solution. Fig. 7 shows a comparison of a typical solution obtained from the computer with the numerical solution. The temperature variation for several positions in the slab, and the percentage of unmelted material are plotted as functions of time. The dotted lines represent Landau's solution. They show good agreement in the early stages of melting, and deviate, as one would expect, as the insulated wall begins to influence the solution. The solution was not carried to complete melting, because this would require an infinite amount of computing time which can be understood by reference to Equation [9]. In practice the solution can be run sufficiently close to complete melting in 1 or 2 min, and the end point can be obtained by extrapolation.

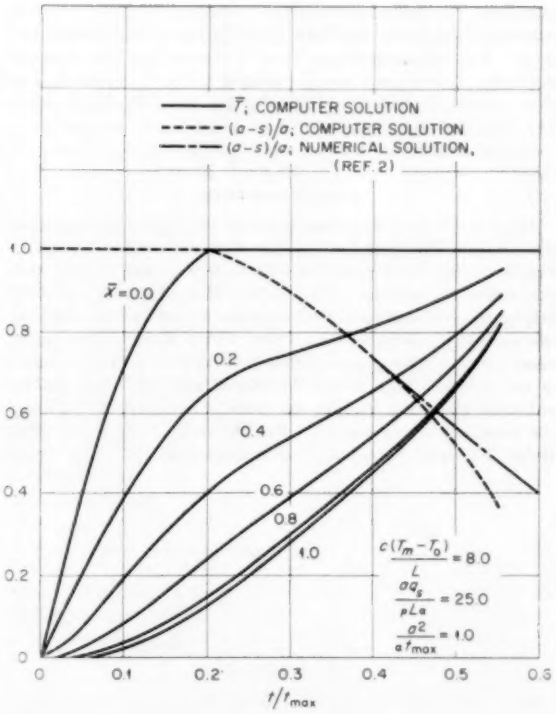


FIG. 7

CONCLUSIONS

A method has been presented for solving the heat-conduction equation with a melting boundary condition. The advantages of this method are (*a*) the melting process is represented continuously with time, and (*b*) a minimum number of cells is required since their number is constant throughout the solution. The circuit has been used with a wide variety of boundary conditions over a period of several months. It has been found to be quite versatile and was relatively trouble-free after an initial period of "debugging." There has been little mention of the

actual results obtained in solving this problem, no presentation of generalized solutions, and little detail of the actual circuit diagram. It is not these things, but it is the method that deserves emphasis; it provides a tool for solving a type of problem that often cannot be solved analytically nor be generalized easily. The details of the circuit depend so greatly upon the particular computer installation that it seemed of little value to include them.

Acknowledgment

There are several who deserve credit for the work reported in this paper. The members of the Convair Analog Computer Facility were always generous with their time and patient with the author's questions. Mr. Robert Horowitz was especially helpful in problems associated with the circuit and in obtaining the necessary computer time. Mr. Victor Kebely contributed many worth-while suggestions that have resulted in improvements of the circuit. Mrs. Gwen Weatherall and Mrs. Kay Baylor patiently made the various calculations necessary to establish the accuracy of the circuit. Dr. Heinz F. Poppendick (Oak Ridge National Laboratory) contributed several very useful criticisms which helped to clarify portions of the text. The author deeply appreciates the contributions made by each one.

BIBLIOGRAPHY

1 "Conduction of Heat in Solids," by H. S. Carslaw and J. C. Jaeger, Oxford University Press, London, England, 1948.

2 "Heat Conduction in a Melting Solid," by H. G. Landau, *Quarterly of Applied Mathematics*, vol. 8, p. 81.

3 "Electric Analogy to Transient Heat Conduction in a Medium With Variable Thermal Properties," by D. R. Otis, Proceedings of the 1956 National Simulation Conference.

4 "The Solution of Partial Differential Equations by Difference Methods Using the Electronic Differential Analyzer," by R. M. Howe and V. S. Haneman, Proceedings of The Institute of Radio Engineers, vol. 41, 1953, pp. 1497–1508.

5 "A Method for Determining Unsteady-State Heat Transfer by Means of an Electric Analogy," by V. Paschkis and H. D. Baker, Trans. ASME, vol. 64, 1942, p. 105.

6 "Transmission Lines and Networks," by W. C. Johnson, McGraw-Hill Book Company, Inc., New York, N. Y., first edition, 1950, p. 25.

7 "Electronic Analog Computers," by G. A. Korn and T. M. Korn, McGraw-Hill Book Company, Inc., New York, N. Y., 1949.

Effect of Axial Fluid Conduction on Heat Transfer in the Entrance Regions of Parallel Plates and Tubes[1]

By P. J. SCHNEIDER,[2] MINNEAPOLIS, MINN.

The laminar heat transfer in thermal-entry regions of parallel plates and tubes is investigated for special conditions under which the contribution of axial fluid conduction is comparable to or greater than the axial convection (low Peclet numbers). Zero and finite axial-conduction solutions for the local and mean fluid temperature and the local total Nusselt number are presented for uniform-velocity flow, finite wall resistance, and for both a uniform ambient temperature and step discontinuity in the ambient temperature. Calculated mean fluid temperatures, local total Nusselt numbers, and thermal-entry lengths are compared with corresponding predictions for the conventional case of zero axial conduction.

Nomenclature

The following nomenclature is used in the paper:

c = uniform thermal capacity of fluid
D = inside tube diameter
E_n = defined by Equation [10]
F_n = defined by Equation [15]
F_n' = defined by Equation [22]
G_n = defined by Equation [26]
Gz = Graetz number (Pe/ξ)
h_a = ambient-side unit surface conductance
J_n = nth-order Bessel function of first kind
k = uniform thermal conductivity of fluid
k_w = uniform thermal conductivity of wall material
Nu = local total Nusselt number ($U\delta/k$ or UD/k)
Nu_a = ambient-side Nusselt number ($U_a\delta/k_w$ or U_aD/k_w)
Pe = Peclet number ($u\delta/\alpha$ or uD/α)
Pr = Prandtl number of fluid ($\rho c\nu/k$)
Re = Reynolds number ($u\delta/\nu$ or uD/ν)
t = local fluid temperature
$\bar{t}$ = local mean fluid temperature
t_a = ambient temperature
t_0 = uniform upstream fluid temperature
u = uniform fluid velocity
U = over-all thermal transmittance
U_a = wall thermal transmittance
x, y = co-ordinates parallel and perpendicular to flow direction
α = thermal diffusivity of fluid ($k/\rho c$)
δ = inside distance between plates
ζ = dimensionless transverse co-ordinate (y/δ or y/D)
η = $\mathrm{Nu}_a/2$
θ = $(t - t_a)/(t_0 - t_a)$
$\bar{\theta}$ = $(\bar{t} - t_a)/(t_0 - t_a)$
μ_n = eigenvalues determined by Equation [11] or [27]
ν = kinematic viscosity of fluid
ξ = dimensionless axial co-ordinate (x/δ or x/D)
ξ_e = thermal-entry length
ρ = mass density of fluid
τ = thickness of plate or tube wall

Subscripts

a = ambient side
e = entry
n = successive positive integers
w = wall
1 = downstream ($\xi > 0$)
2 = upstream ($\xi < 0$)
∞ = fully developed

Introduction

The original Graetz problem (1)[3] of low-speed, hydrodynamically developed, laminar heat transfer to a uniform-property fluid in the thermal-entry region of a uniform wall-temperature tube has, in the intervening years, been extended in a variety of ways. The corresponding problem of steady flow between flat parallel plates has been studied (2), and consideration has been given to such additional effects as wall and external surface resistance (3, 4), nonuniform inlet temperature (5), nonsymmetrical boundary conditions (6), uniform heat-input boundary conditions (7), nonuniform boundary conditions (8), power velocity profiles (9), rectangular duct geometry (10), hydrodynamically undeveloped flow (11, 12), dissipative flow (13, 14), nonuniform properties (14, 15), and so on.

One of the fundamental assumptions on which the solution of the Graetz problem and all its extensions rest, however, is the neglect of fluid conduction in the axial (longitudinal) direction of flow. While this assumption of zero axial thermal conductivity is clearly permissible in many practical cases of nonisothermal conduit flow, it can lead to a significant error for the special case of a low Reynolds-number flow of a low Prandtl-number fluid. It is the purpose of the present study to direct attention to the role of axial fluid conduction under these special conditions for flow between infinite parallel plates and in long tubes. The results should therefore bear on the theory of liquid-metal heat transfer (low Prandtl number) and, in general, have practical significance for conduit flow characterized by a low Peclet number (product of the Reynolds and Prandtl numbers) in which an accurate estimate of the entry-region heat transfer and the thermal-entry length is required.

The over-all effects of axial conduction are concentrated in the thermal-entry region of the conduit, and as such it does not influence the local heat transfer in the downstream region of thermally established flow. In other words the local asymptotic Nusselt numbers are the same either in the presence of or in the ab-

[1] A publication of the Heat Transfer Laboratory, University of Minnesota, Minneapolis, Minn.

[2] Research Associate, Department of Mechanical Engineering, University of Minnesota.

Contributed by the Heat Transfer Division of The American Society of Mechanical Engineers and presented at the Heat Transfer and Fluid Mechanics Institute, Stanford University, Calif., June 21–23, 1956.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, May 11, 1956.

[3] Numbers in parentheses refer to Bibliography at end of paper.

sence of axial fluid conduction. But since the axial conduction does affect the local heat transfer in the entrance region, the integrated or average heat exchange is distinctly different if one takes this effect into account.

Now the possibility of an axial-conduction thermal-entry effect occurring at low Peclet numbers is apparent as soon as one recognizes that without axial conduction the fluid temperature approaches the wall or ambient temperature as the Peclet number tends to zero. But in the presence of axial conduction the fluid and ambient temperatures are evidently not the same for the pure-conduction case of zero Peclet number.

The approximate order of magnitude of the Peclet number at which this conduction effect becomes important can be estimated from the relative magnitudes of the longitudinal conduction and convection terms (3). The axial conduction is $k\partial^2 t/\partial x^2$ and the axial convection is $\rho cu\partial t/\partial x$, where k, ρ, and c represent the uniform thermal conductivity, mass density, and thermal capacity of the fluid, u the uniform flow velocity, t the fluid temperature, and x the axial co-ordinate in the direction of flow. The orders of magnitude of these two terms are k/L^2 and $\rho cu/L$, where L is a characteristic order-of-magnitude length such as plate spacing or tube diameter, and hence their ratio is of the order

$$\text{Axial conduction/axial convection} = 1/\text{Pe}$$

This suggests that the effect of axial conduction is generally negligible for Peclet numbers exceeding, say, 100. The axial-conduction effect therefore should be examined in the approximate range

$$0 \leq \text{Re Pr} \leq 100$$

An analysis of this problem for a fully developed parabolic velocity profile (Poiseuille flow) leads to a Whittaker differential equation and a general solution for which the eigenfunctions are not orthogonal (6). Hence the constants for the series solution cannot be obtained in the usual analytical way. On the other hand, if one assumes a uniform velocity profile (slug flow), then the eigenfunctions are orthogonal and many of the required eigenvalues are already known from solutions within the field of pure conduction. It therefore appears that until these eigenvalues for the general problem with parabolic velocity become available, the results for slug flow may be useful in indicating the relative magnitude of the conduction effects to anticipate in the case of Poiseuille flow.[4]

The assumption of uniform velocity yields an upper limit for the calculated heat transfer. This assumption may be a reasonable one if the Prandtl number is sufficiently low, since under these conditions the entry-region temperature profile develops at a more rapid rate than the velocity profile. This suggests, in turn, the possibility of a low Peclet number for which the effect of axial fluid conduction is important. The results for uniform velocity also may indicate the gross effects to be expected in the case of turbulent flow at low Peclet numbers.

Differential Equations and Boundary Conditions

The various cases analyzed are classified for both parallel plates and tubes according as to whether the effect of axial conduction is excluded or included, and according to the boundary conditions (abbreviated BC) assumed:

Case I—Without axial conduction
Case II—With axial conduction
BC.1—Uniform ambient temperature
BC.2—Uniform step ambient temperature

[4] Approximate solutions for Poiseuille flow can be obtained readily by numerical means (11), since by this method the shape of the velocity profile is relatively insignificant in so far as it affects the labor of the numerical solution.

Let the longitudinal axis of a pair of infinite parallel plates or a tube be denoted by x and the transverse or radial axis be denoted by y. The orientation of the co-ordinate system is shown in Fig. 1 with dimensionless co-ordinates $\xi = x/\delta = x/D$ and $\zeta = y/\delta = y/D$ where δ is the inside transverse distance between the plates

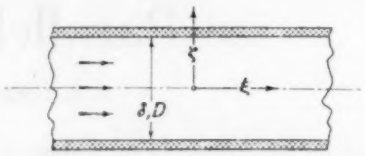


Fig. 1 Dimensionless Co-Ordinates ξ and ζ for Parallel Plates and Tubes

and D is the inside diameter of the tube. In addition let a dimensionless excess temperature be defined as

$$\theta(\xi, \zeta) = (t - t_a)/(t_0 - t_a)$$

where $t = t(\xi, \zeta)$ is the nonuniform temperature of the fluid, t_a is the uniform ambient temperature, and t_0 is the uniform temperature of the upstream fluid. The flow is from left to right, and is assumed to be fully established (transverse or radial components of velocity are zero) and uniform across the duct as u. The properties of the fluid are assumed to be uniform, and conversion of mechanical into thermal energy is considered negligible since only small Peclet numbers (low velocities) are considered. Under these conditions, the partial-differential equations to be satisfied by θ for steady nondissipative flow between parallel plates and in tubes are, respectively

$$\frac{\partial^2\theta}{\partial\xi^2} + \frac{\partial^2\theta}{\partial\zeta^2} - \text{Pe}\,\frac{\partial\theta}{\partial\xi} = 0 \qquad [1]$$

and

$$\frac{\partial^2\theta}{\partial\xi^2} + \frac{\partial^2\theta}{\partial\zeta^2} + \frac{1}{\zeta}\frac{\partial\theta}{\partial\zeta} - \text{Pe}\,\frac{\partial\theta}{\partial\xi} = 0 \qquad [2]$$

where the dimensionless Peclet number is given by $\text{Pe} = \text{Re Pr} = u\delta/\alpha = uD/\alpha$, u being the uniform velocity of flow and α the thermal diffusivity ($k/\rho c$) of the fluid. In these equations the first term represents axial conduction, the middle term(s) account for transverse conduction, and the last term represents axial convection. For Case II the axial-conduction term is present, while for Case I $\partial^2\theta/\partial\xi^2 = 0$.

The first set of boundary conditions considered is, for both infinite parallel plates and long tubes

BC.1

$$\left.\begin{aligned} &(a)\ \ \theta(0, \zeta) = 1 \\ &(b)\ \ \theta\left(\xi, \frac{1}{2}\right) = -\frac{1}{\text{Nu}_a}\left[\frac{\partial}{\partial\zeta}\,\theta(\xi, \zeta)\right]_{\xi,\,1/2} \\ &(c)\ \ \theta(\infty, \zeta) \to 0 \\ &(d)\ \ \frac{\partial}{\partial\zeta}\,\theta(\xi, 0) = 0 \end{aligned}\right\} \qquad [3]$$

The first condition (a) suggests that the heating or cooling begins at the axial station $\xi = 0$, where the fluid temperature is a uniform t_0. Condition (b) describes the heat loss (or gain) through the duct walls. The uniform conductance of the walls is given by a wall or "ambient-side" Nusselt number $\text{Nu}_a = U_w\delta/k = U_wD/k$, where k is the thermal conductivity of the fluid and U_w is a wall thermal transmittance which combines the resistance of the wall and the ambient-side surface resistance. Con-

sidering a plate thickness of τ and a uniform unit surface conductance on the ambient side as h_a, then $U_w = (\tau/k_w + 1/h_a)^{-1}$. If Nu_a is finite, then a temperature rise (or drop) occurs at the external surface and/or through the wall, while if $Nu_a = \infty$ the wall and external surface resistances are zero which means that the ambient and wall temperatures are uniform and equal. For the completely insulated case in which either the wall or external surface resistance is infinite, $Nu_a = 0$. The third condition (c) fixes the asymptotic behavior of θ ($t \to t_a$ at large ξ) and the fourth condition (d) implies temperature symmetry about $\zeta = 0$.

The second set of boundary conditions considered are, for both parallel plates and tubes

$$\left.\begin{aligned}
&(a)\ \theta_1(0, \zeta) = \theta_2(0, \zeta)\\
&(b)\ (\partial\theta_1/\partial\xi)_{0,\zeta} = (\partial\theta_2/\partial\xi)_{0,\zeta}\\
&(c)\ \theta_1\left(\xi, \frac{1}{2}\right) = -\frac{1}{Nu_a}\left[\frac{\partial}{\partial\zeta}\,\theta_1(\xi, \zeta)\right]_{\xi,\,1/2}\\
&(d)\ \theta_1(\infty, \zeta) \to 0\\
&\text{BC.2}\\
&(e)\ \frac{\partial}{\partial\zeta}\,\theta_2(\xi, 0) = 0\\
&(f)\ \theta_2\left(-\xi, \frac{1}{2}\right) = 1 - \frac{1}{Nu_a}\left[\frac{\partial}{\partial\zeta}\,\theta_2(\xi, \zeta)\right]_{-\xi,\,1/2}\\
&(g)\ \theta_2(-\infty, \zeta) \to 1\\
&(h)\ \frac{\partial}{\partial\zeta}\,\theta_2(-\xi, 0) = 0
\end{aligned}\right\} \quad \ldots [4]$$

where $\theta_1 = \theta(\xi, \zeta)$ to the right (downstream) of the origin $\xi = 0$, and $\theta_2 = \theta(-\xi, \zeta)$ to the left (upstream) of the origin. The first and second conditions (a) and (b) express continuity of the temperature and heat flux for the two solutions θ_1 and θ_2 at the plane $\xi = 0$. Boundary conditions (c) and (f) are the same as in BC.1, where now the ambient-side Nusselt number Nu_a is assumed to be the same for both $\xi > 0$ and $\xi < 0$. The asymptotic and symmetry conditions (d), (e), (g), and (h) are the same as in BC.1.

Analytical Results

Solutions for Cases I and II are presented here for the local fluid temperature $\theta(\xi, \zeta)$, mean fluid temperature, $\bar{\theta}(\xi)$, and local total Nusselt number, $Nu(\xi)$, for parallel plates and tubes with BC.1 and BC.2. For each of the solutions the mean fluid temperature is calculated as

$$\bar{\theta} = 2\int_0^{1/2} \theta\, d\zeta \quad \text{and} \quad \bar{\theta} = 8\int_0^{1/2} \theta\zeta\, d\zeta \ldots . [5]$$

for parallel plates and tubes, respectively, and the local total Nusselt number which determines the local heat exchange between the fluid and duct environment is calculated according to

$$Nu = -\frac{1}{\bar{\theta}}\left(\frac{\partial\theta}{\partial\zeta}\right)_{1/2} \ldots\ldots [6]$$

In Equation [6] $Nu = U\delta/k$ or UD/k, where U is the over-all thermal transmittance between the fluid and external duct environment.

Parallel Plates

1 Case I, BC.1. The particular solution satisfying Equation [1] with $\partial^2\theta/\partial\xi^2 = 0$ and the boundary conditions in Equation [3] is found to be

$$\theta = 2\sum_0^\infty \frac{E_n\mu_n}{\sin\mu_n}\, e^{-4\mu_n^2\xi/Pe}\cos 2\mu_n\zeta \ldots\ldots [7]$$

so that, by Equations [5] and [6]

$$\bar{\theta} = 2\sum_0^\infty E_n e^{-4\mu_n^2\xi/Pe} \ldots\ldots [8]$$

and

$$Nu = \frac{4}{\bar{\theta}}\sum_0^\infty E_n\mu_n^2 e^{-4\mu_n^2\xi/Pe} \ldots\ldots [9]$$

In Equations [7], [8], and [9] the E_n are abbreviations for

$$E_n = \frac{1}{\mu_n}\left(\frac{\sin^2\mu_n}{\mu_n + \sin\mu_n\cos\mu_n}\right) \ldots\ldots [10]$$

and the μ_n are eigenvalues satisfying

$$\mu_n \tan\mu_n = \eta;\ n = 0, 1, 2, \ldots\ldots [11]$$

The first six roots (16) of Equation [11] are given in Table 1 for seven values of $\eta = Nu_a/2$.

Results for the special case of zero wall resistance also are contained in Equations [7], [8], and [9], but will not be explicitly written down here. For that case $\eta = \infty$ and the roots of Equation [11] are simply $\mu_n = (2n + 1)\pi/2$, $n = 0, 1, 2, \ldots$ (Table 1), and $E_n = 1/\mu_n^2$.

The solutions for Case I are all expressed in terms of a single parameter Pe/ξ, which is the familiar dimensionless Graetz number $Gz = Re\,Pr/\xi$. Such is not possible for the following solutions of Case II in which the influence of axial-fluid conduction is taken into account.

2 Case II, BC.1. The corresponding solution with $\partial^2\theta/\partial\xi^2 \neq 0$ is

$$\theta = 2\sum_0^\infty \frac{E_n\mu_n}{\sin\mu_n}\, e^{-F_n\xi}\cos 2\mu_n\zeta \ldots\ldots [12]$$

whereby

$$\bar{\theta} = 2\sum_0^\infty E_n e^{-F_n\xi} \ldots\ldots [13]$$

and

$$Nu = \frac{4}{\bar{\theta}}\sum_0^\infty E_n\mu_n^2 e^{-F_n\xi} \ldots\ldots [14]$$

In Equations [12], [13], and [14] the E_n and μ_n are again given in Equation [10] and Table 1, and the F_n are defined as

$$F_n = [(Pe/2)^2 + 4\mu_n^2]^{1/2} - Pe/2 \ldots\ldots [15]$$

Thus, the solutions including axial fluid conduction cannot be expressed in terms of the usual Graetz number Gz. Again, Equa-

TABLE 1 FIRST SIX ROOTS μ_n OF EQUATION [11]

η	μ_0	μ_1	μ_2	μ_3	μ_4	μ_5
0.25	0.4773	3.2190	6.3226	9.4512	12.5863	15.7239
0.50	0.6533	3.2923	6.3616	9.4775	12.6060	15.7397
1	0.8603	3.4256	6.4373	9.5293	12.6453	15.7713
5	1.3138	4.0336	6.9096	9.8928	12.9352	16.0107
10	1.4289	4.3058	7.2281	10.2003	13.2142	16.2594
50	1.5400	4.6202	7.7012	10.7832	13.8666	16.9519
∞	$\pi/2$	$3\pi/2$	$5\pi/2$	$7\pi/2$	$9\pi/2$	$11\pi/2$

tions [12], [13], and [14] include as a special case the results for $Nu_a = \infty$. Also note by Equations [12] and [15] that for Case II, $\theta \neq 0$ for $Pe = 0$. Thus, these and all subsequent results for Case II contain the well-known pure-conduction solutions (17) for which $Pe = 0$.

3 Case I, BC.2. By matching solutions θ_1 and θ_2 satisfying Equation [1] with $\partial\theta/\partial\xi^2 = 0$ and the boundary conditions in Equation [4], one finds

$$\left.\begin{aligned} \xi > 0;\ \theta_1 &= \sum_0^\infty \frac{E_n\mu_n}{\sin\mu_n} e^{-4\mu_n^2\xi/Pe}\cos 2\mu_n\zeta \\ \xi < 0;\ \theta_2 &= 1 - \sum_0^\infty \frac{E_n\mu_n}{\sin\mu_n} e^{4\mu_n^2\xi/Pe}\cos 2\mu_n\zeta \end{aligned}\right\} \quad ..[16]$$

Thus, for the downstream region

$$\xi > 0; \quad \bar\theta_1 = \sum_0^\infty E_n e^{-4\mu_n^2\xi/Pe} \quad .....[17]$$

which is one half that given by Equation [8] for BC.1, and

$$\xi > 0;\ Nu_1 = \frac{2}{\bar\theta_1}\sum_0^\infty E_n\mu_n^2 e^{-4\mu_n^2\xi/Pe} \quad .....[18]$$

which is the same as Equation [9] for BC.1.

4 Case II, BC.2.[5] The corresponding solutions with $\partial^2\theta/\partial\xi^2 \neq 0$ are

$$\left.\begin{aligned} \xi > 0;\ \theta_1 &= 2\sum_0^\infty \frac{E_n\mu_n}{\sin\mu_n}\frac{e^{-F_n\xi}}{(F_n/F_n' + 1)}\cos 2\mu_n\zeta \\ \xi < 0;\ \theta_2 &= 1 - 2\sum_0^\infty \frac{E_n\mu_n}{\sin\mu_n}\frac{e^{F_n'\xi}}{(F_n'/F_n + 1)}\cos 2\mu_n\zeta \end{aligned}\right\} \quad ..[19]$$

whereby

$$\xi > 0;\ \bar\theta_1 = 2\sum_0^\infty \frac{E_n}{(F_n/F_n' + 1)} e^{-F_n\xi} \quad .....[20]$$

and

$$\xi > 0; \quad Nu_1 = \frac{4}{\bar\theta_1}\sum_0^\infty \frac{E_n\mu_n^2}{(F_n/F_n' + 1)} e^{-F_n\xi} \quad .....[21]$$

In Equations [19], [20], and [21] the F_n are given by Equation [15] and the F_n' are defined as

$$F_n' = [(Pe/2)^2 + 4\mu_n^2]^{1/2} + Pe/2 \quad .....[22]$$

Tubes

5 Case I, BC.1. The particular solution satisfying Equation [2] with $\partial^2\theta/\partial\xi^2 = 0$ and the boundary conditions in Equation [3] is found to be

$$\theta = 2\sum_1^\infty \frac{G_n\mu_n}{J_1(\mu_n)} e^{-4\mu_n^2\xi/Pe} J_0(2\mu_n\zeta) \quad .....[23]$$

so that by Equations [5] and [6]

$$\bar\theta = 4\sum_1^\infty G_n e^{-4\mu_n^2\xi/Pe} \quad .....[24]$$

and

$$Nu = \frac{4}{\bar\theta}\sum_1^\infty G_n\mu_n^2 e^{-4\mu_n^2\xi/Pe} \quad .....[25]$$

In Equations [23], [24], and [25] the G_n are abbreviations for

$$G_n = 1/\mu_n^2[J_0^2(\mu_n)/J_1^2(\mu_n) + 1] \quad .....[26]$$

and the μ_n are now eigenvalues satisfying

$$\mu_n J_1(\mu_n)/J_0(\mu_n) = \eta; \quad n = 1, 2, 3, \ldots \quad .....[27]$$

The first six roots (16) of Equation [27] are given in Table 2 for seven values of $\eta = Nu_a/2$.

For the special case of $Nu_a = \infty$, the eigenvalues μ_n in Equation [27] become the consecutive roots of $J_0(\mu_n) = 0$, $n = 1,2,3,\ldots$ (Table 2), and $G_n = 1/\mu_n^2$.

6 Case II, BC.1. The corresponding solution with $\partial^2\theta/\partial\xi^2 \neq 0$ is

$$\theta = 2\sum_1^\infty \frac{G_n\mu_n}{J_1(\mu_n)} e^{-F_n\xi} J_0(2\mu_n\zeta) \quad .....[28]$$

whereby

$$\bar\theta = 4\sum_1^\infty G_n e^{-F_n\xi} \quad .....[29]$$

and

$$Nu = \frac{4}{\bar\theta}\sum_1^\infty G_n\mu_n^2 e^{-F_n\xi} \quad .....[30]$$

In Equations [28], [29], and [30], the F_n, G_n, and μ_n are again given in Equations [15] and [26] and in Table 2.

7 Case I, BC.2. The corresponding solutions θ_1 and θ_2 satisfying Equation [2] with $\partial^2\theta/\partial\xi^2 = 0$ and the boundary conditions in Equation [4] are

$$\left.\begin{aligned} \xi > 0; \quad \theta_1 &= \sum_1^\infty \frac{G_n\mu_n}{J_1(\mu_n)} e^{-4\mu_n^2\xi/Pe} J_0(2\mu_n\zeta) \\ \xi < 0; \quad \theta_2 &= 1 - \sum_1^\infty \frac{G_n\mu_n}{J_1(\mu_n)} e^{4\mu_n^2\xi/Pe} J_0(2\mu_n\zeta) \end{aligned}\right\} \quad ..[31]$$

Thus, for the downstream region

$$\xi > 0; \quad \bar\theta_1 = 2\sum_1^\infty G_n e^{-4\mu_n^2\xi/Pe} \quad .....[32]$$

and

$$\xi > 0; \quad Nu_1 = \frac{2}{\bar\theta_1}\sum_1^\infty G_n\mu_n^2 e^{-4\mu_n^2\xi/Pe} \quad .....[33]$$

[5] A solution for this case with $Nu_a = \infty$ was originally obtained by Wilson (18).

TABLE 2 FIRST SIX ROOTS μ_n OF EQUATION [27]

η	μ_1	μ_2	μ_3	μ_4	μ_5	μ_6
0.25	0.6818	3.8063	7.0511	10.1980	13.3425	16.4858
0.50	0.9408	3.9594	7.0864	10.2225	13.3611	16.5010
1	1.2558	4.0795	7.1558	10.2710	13.3984	16.5312
5	1.9890	4.7131	7.6177	10.6223	13.6786	16.7630
10	2.1795	5.0332	7.9569	10.9363	13.9580	17.0099
50	2.3572	5.4112	8.4840	11.5621	14.6433	17.7272
∞	2.4048	5.5201	8.6537	11.7915	14.9309	18.0711

8 Case II, BC.2. The solutions with $\partial^2\theta/\partial\xi^2 \neq 0$ are

$$\left.\begin{aligned} \xi > 0;\ \theta_1 &= 2\sum_1^\infty \frac{G_n\mu_n}{J_1(\mu_n)}\,\frac{e^{-F_n\xi}}{(F_n/F_n'+1)}\,J_0(2\mu_n\zeta) \\ \xi > 0;\ \theta_2 &= 1 - 2\sum_1^\infty \frac{G_n\mu_n}{J_1(\mu_n)}\,\frac{e^{F_n'\xi}}{(F_n'/F_n+1)}\,J_0(2\mu_n\zeta) \end{aligned}\right\} \quad ..[34]$$

whereby

$$\xi > 0; \quad \bar\theta_1 = 4\sum_1^\infty \frac{G_n}{(F_n/F_n'+1)}\,e^{-F_n\xi} \quad .....[35]$$

and

$$\xi > 0; \quad \mathrm{Nu}_1 = \frac{4}{\bar\theta_1}\sum_1^\infty \frac{G_n\mu_n^2}{(F_n/F_n'+1)}\,e^{-F_n\xi} \quad .....[36]$$

in which the F_n' are again given in Equation [22].

Results of Calculation

Mean Temperature. Some typical temperature fields in a tube with axial fluid conduction (Case II) are shown in Fig. 2 for Pe = 1. The first (upper) set of isotherms was calculated by Equation [28] for zero resistance [uniform wall temperature, $J_0(\mu_n) = 0$]. The second set of isotherms, Equation [28], shows the influence of wall and external surface resistance (uniform ambient temperature) with $\mathrm{Nu}_a = 1$. The third temperature field was calculated by Equations [34] for zero resistance [step wall temperature, $J_0(\mu_n) = 0$], and the last (lower) set of isotherms, Equations [34], illustrates the effect of wall and surface resistance (step ambient temperature) with $\mathrm{Nu}_a = 1$. Under this set of conditions, the effect of wall resistance on the temperature conformation in the fluid is somewhat more pronounced for BC.1 than for BC.2.

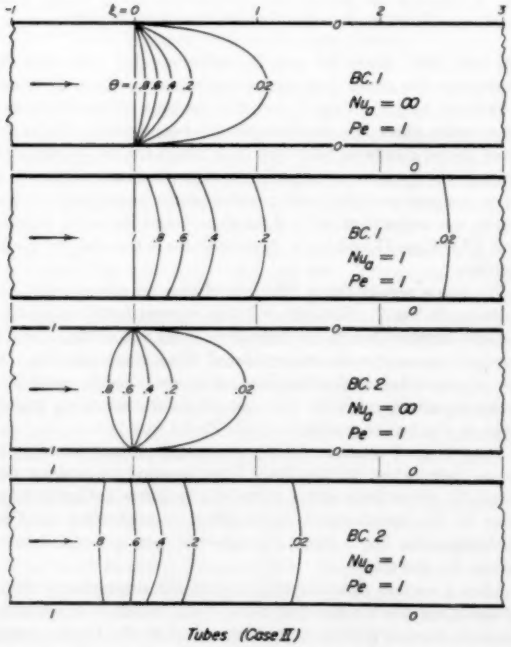


Fig. 2 Temperature Fields θ in a Tube With Axial Fluid Conduction for BC.1 and BC.2

A comparison of mean fluid temperatures $\bar\theta$ for BC.1 and Cases I and II is shown in Fig. 3 for parallel plates, Equations [8] and [13], and in Fig. 4 for tubes, Equations [24] and [29]. The results suggest that neglecting axial fluid conduction tends to underestimate the mean fluid temperature at all axial stations ξ. The error becomes more pronounced as the Peclet number decreases, and is a maximum for the pure-conduction case (Pe = 0) in which there is no contribution to the heat transfer by axial convection. In that case $\bar\theta$ for Case II and Pe = 0 is even well above $\bar\theta$ for Case I and Pe = 1. This effect, on the other hand, is almost completely absent for both parallel plates and tubes for Re Pr ≥

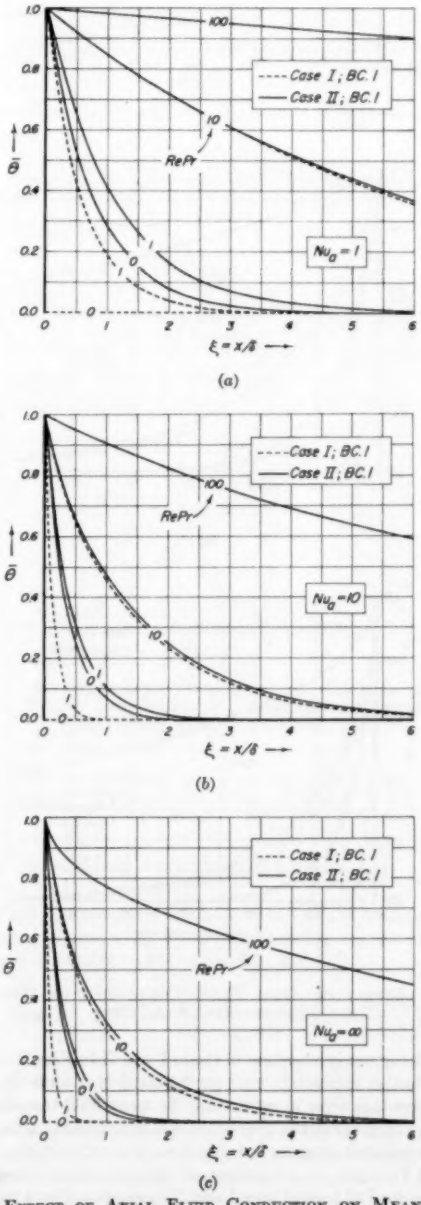
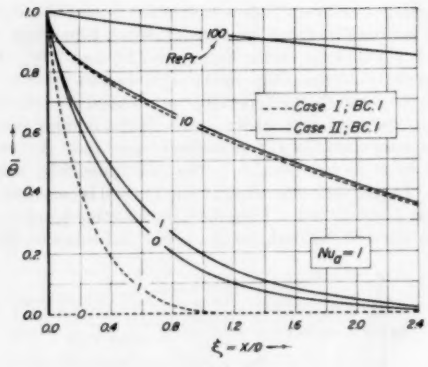
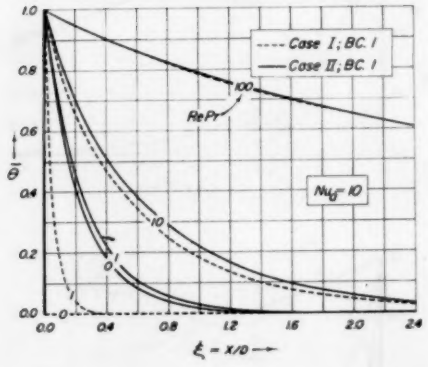


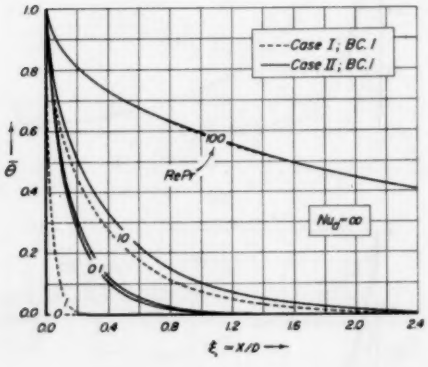
Fig. 3 Effect of Axial Fluid Conduction on Mean Fluid Temperature $\bar\theta$ Between Parallel Plates



(a)

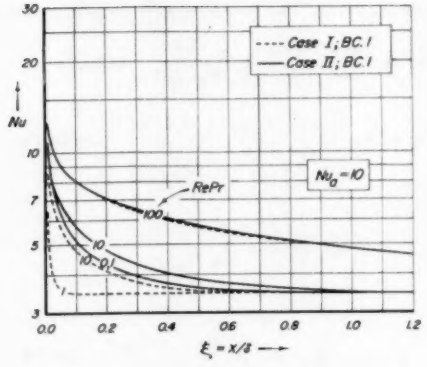


(b)

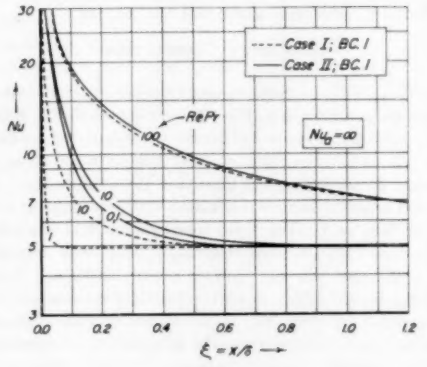


(c)

FIG. 4 EFFECT OF AXIAL FLUID CONDUCTION ON MEAN FLUID TEMPERATURE $\bar{\theta}$ IN TUBES



(a)



(b)

FIG. 5 EFFECT OF AXIAL FLUID CONDUCTION ON LOCAL TOTAL NUSSELT NUMBER Nu FOR PARALLEL PLATES

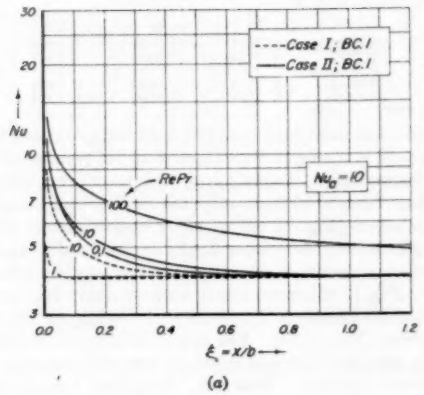
100. A further comparison of Cases I and II in Figs. 3 and 4 suggests that an increase in wall conductance Nu_a tends to increase the percentage error in estimating the mean fluid temperature by Case I. In general, the over-all effects of axial conduction appear to be somewhat stronger in tubes than in parallel plates.

Heat Transfer. A comparison of calculated local Nusselt numbers Nu for BC.1 and Cases I and II is given in Fig. 5 for parallel plates, Equations [9] and [14], and in Fig. 6 for tubes, Equations [25] and [30]. Since for zero Reynolds number and zero axial conduction the mean fluid temperature is zero, the heat transfer is likewise zero for Case I and this limiting value of Pe = 0. Here, again, there is a considerable axial-conduction effect at the lower Peclet numbers, and very little effect for Pe $\geq$ 100. This effect shows up more strongly in the Nusselt numbers than in the mean temperatures for both parallel plates and tubes. For example, one notes that while $\bar{\theta}$ for Case I and Pe = 10 is greater than $\bar{\theta}$ for Case II and Pe = 1, the reverse is true for the Nusselt numbers.

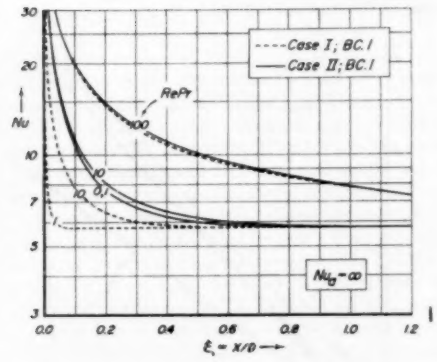
The same results for a tube are shown on more familiar coordinates in Fig. 7. Here the ordinate represents the ratio of local Nusselt number Nu to its fully developed value Nu_∞, and the abscissa represents the conventional Graetz number Gz. The case of zero axial conduction gives, of course, a single curve for all Peclet numbers, while in the case of axial conduction included there is a separate curve for each Peclet number.

Mean Nusselt numbers for a given length of duct are evaluated by an integration of the local Nusselt numbers over a finite length ξ. Since these mean values will be lower in Case I than in Case II, one concludes that the effect of neglecting axial fluid conduction for Pe < 100 is to underestimate the total heat exchange for the duct.

After a certain thermal-entry length the downstream Nusselt numbers approach a limiting value Nu_∞, which is dependent on the wall Nusselt number but independent of the Peclet number. These fully developed Nusselt numbers are those for which the second and succeeding terms in the series for Nu are negligibly



(a)



(b)

FIG. 6 EFFECT OF AXIAL FLUID CONDUCTION ON LOCAL TOTAL NUSSELT NUMBER Nu FOR TUBES

small. Since axial conduction affects only the exponential terms in the series, it is apparent that these asymptotic values will be the same for both Cases I and II. Thus, by Equations [8] and [9] for parallel plates $Nu_\infty = 2\mu_0^2$, and by Equations [24] and [25] for tubes $Nu_\infty = \mu_1^2$. Referring to Tables 1 and 2 one finds for plates (Fig. 5) $Nu_\infty = 3.452$ and 4.935 for $Nu_0 = 10$ and ∞, and for tubes (Fig. 6) $Nu_\infty = 3.959$ and 5.783 for $Nu_0 = 10$ and ∞. A large range of asymptotic Nusselt numbers is given in Fig. 8 for interpolation between $Nu_0 = 0.1$ and 100. The curves approach each other at the lower values of the wall Nusselt number Nu_0, since under these conditions the influence of duct shape disappears and the heat transfer is predominantly controlled by the resistance of the duct wall. Also shown for comparison in Fig. 8 are the corresponding asymptotic Nusselt numbers for a fully developed parabolic velocity profile (3). Although the growth of these asymptotic values with decreasing resistance is analogous for both velocity distributions, the influence of a parabolic velocity is seen to decrease the predicted values of Nu_∞ and to diminish the influence of duct shape. The effect of duct shape is also seen to be opposite for the two velocity profiles.

Thermal-Entry Length. As a criterion for estimating the thermal-entry length ξ_e (value of ξ for which the Nusselt number is fully developed as Nu_∞), Berry (19) has suggested that in the thermally established flow the second term in the series for Nu should not contribute more than 1 per cent of the first. An interpretation of this in terms of Equations [9] and [14] for parallel plates (BC 1) gives

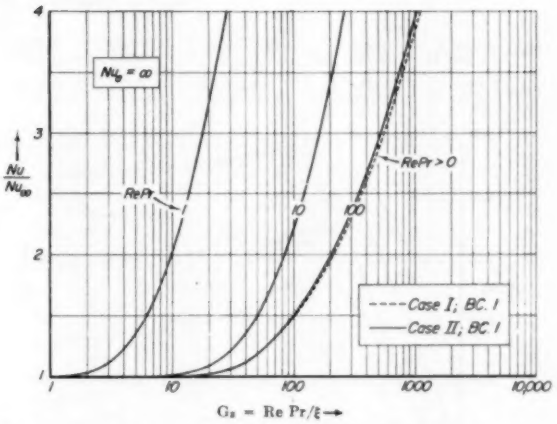


FIG. 7 RATIO OF LOCAL TOTAL NUSSELT NUMBER Nu TO ITS FULLY DEVELOPED VALUE Nu_∞ IN A TUBE AS A FUNCTION OF GRAETZ NUMBER Gz

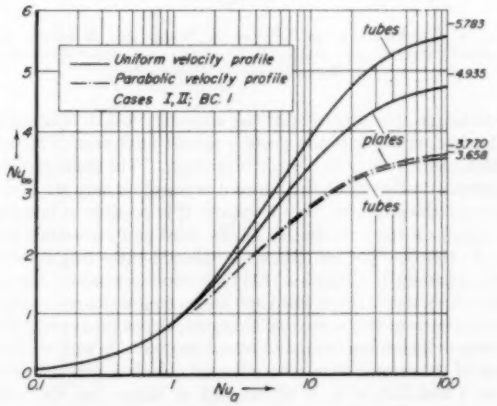


FIG. 8 FULLY DEVELOPED LOCAL NUSSELT NUMBERS Nu_∞ FOR PARALLEL PLATES AND TUBES AS A FUNCTION OF WALL NUSSELT NUMBER Nu_0, AND VELOCITY PROFILE

$$\text{Case I:}\quad \xi_e/\text{Pe} = \frac{1}{4(\mu_1^2 - \mu_0^2)} \ln\left(100\,\frac{\mu_1^2}{\mu_0^2}\frac{E_1}{E_0}\right)$$

$$\text{Case II:}\quad \xi_e = \frac{1}{(F_1 - F_0)} \ln\left(100\,\frac{\mu_1^2}{\mu_0^2}\frac{E_1}{E_0}\right)$$

in which the μ_n's are given in Table 1, and the E_n's and F_n's by Equations [10] and [15]. The corresponding results for tubes (BC.1) are, according to Equations [25] and [30]

$$\text{Case I:}\quad \xi_e/\text{Pe} = \frac{1}{4(\mu_2^2 - \mu_1^2)} \ln\left(100\,\frac{\mu_2^2}{\mu_1^2}\frac{G_2}{G_1}\right)$$

$$\text{Case II:}\quad \xi_e = \frac{1}{(F_2 - F_1)} \ln\left(100\,\frac{\mu_2^2}{\mu_1^2}\frac{G_2}{G_1}\right)$$

in which the μ_n's are given in Table 2 and the G_n's by Equation [26]. Thus, only in Case I is the thermal-entry length proportional to the Peclet number; in this case one defines a reciprocal thermal-entry Graetz number $Gz_e = \text{Pe}/\xi_e$ which is a function of the wall Nusselt number Nu_0 only.

TABLE 3 THERMAL-ENTRY LENGTHS ξ_e FOR BC.1

		Nu_a	0.5	1	2	4	10	20	60	100	∞
Plate	Case I	ξ_e/Pe	0.0404	0.0556	0.0681	0.0746	0.0727	0.0676	0.0620	0.0606	0.0584
	Case II	ξ_e (Pe = 0)	0.2982	0.4390	0.5840	0.7042	0.7772	0.7758	0.7539	0.7463	0.7329
Tube	Case I	ξ_e/Pe	0.0213	0.0331	0.0446	0.0532	0.0564	0.0538	0.0495	0.0484	0.0466
	Case II	ξ_e (Pe = 0)	0.1946	0.3243	0.4755	0.6273	0.7557	0.7754	0.7596	0.7524	0.7391

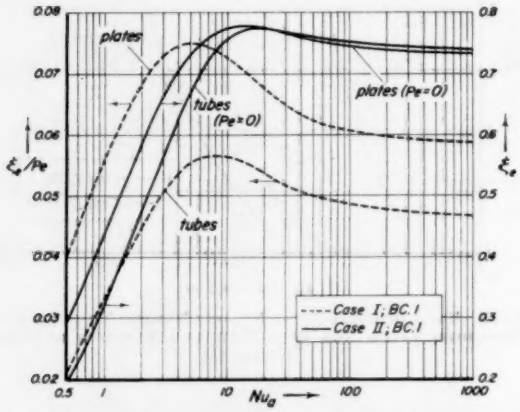


FIG. 9 THERMAL-ENTRY LENGTH ξ_e WITH AND WITHOUT AXIAL FLUID CONDUCTION IN PARALLEL PLATES AND TUBES AS A FUNCTION OF WALL NUSSELT NUMBER Nu_a

Although this definition of thermal-entry length is admittedly arbitrary, it nevertheless gives a simple analytical relationship which yields results in correct magnitude. The definition is also mathematically unjustified, since it is based on only the first two terms of the series for Nu. However, it is possible to investigate in each case the contribution of the third and succeeding terms. To do this one first calculates the thermal-entry length for each Nu_a, assuming by definition that the second term is only 1 per cent of the first and that the third and succeeding terms are negligible. On the basis of this estimated entry length one then calculates the third and remaining terms and compares their sum with the magnitude of the second term. An informal calculation of this sort for Case I and Nu_a = 1, 4, 10, 60, and ∞ shows that the average contribution of this sum for parallel plates is only 0.01 per cent of the second term. A recalculation of these sums based on a longer or shorter entry length decreases or increases their contribution; however, in each case one finds that the chosen entry length fails to satisfy the modified definition that the sum of the second and succeeding terms is 1 per cent of the first. Hence, in the absence of a less arbitrary and more rigorous description of the thermal-entry length, the present definition is considered adequate.

Some results of calculation are given in Table 3 and in Fig. 9. The thermal-entry length in Case I is seen to be considerably shorter for tubes than for parallel plates at all values of Nu_a. The effect of increasing wall conductance in Case I for both parallel plates and tubes is to first increase the thermal-entry length to a maximum value; beyond this the entry length decreases slightly to a limiting value for zero wall resistance. This behavior can be seen by close inspection of Figs. 5 and 6. Also shown in Fig. 9 are the thermal-entry lengths for Case II and Pe = 0. The same general trends persist in the presence of axial conduction except that now the entry lengths for the parallel plates exceed those for the tube only up to their maximum values, while beyond these the plate entry lengths decrease to a slightly lower limit than that for the tube. However, for Pe = 1 the asymptotic value of the plates and tube are nearly identical, and for Pe > 1 the limiting entry lengths for the plate all exceed those for the tube as in Case I.

The effect of axial fluid conduction on the calculated thermal-entry length is shown in Figs. 10 and 11 for parallel plates and tubes, respectively. According to these results, neglecting axial conduction tends to underestimate the thermal-entry length at all Peclet numbers (Figs. 5 and 6) below approximately 100. The error increases with decreasing Peclet number and is, as expected from previous results, a maximum for the pure-conduction case Pe = 0. Fig. 12 illustrates this in terms of a flow Reynolds number Re and the particular fluids; water (Pr ≈ 7.3), air (Pr ≈ 0.73), and mercury (Pr ≈ 0.01). The largest error in the thermal-entry length is near the inlet and is always less than one plate spacing or one tube diameter. This error disappears completely about three or four tube diameters downstream depending somewhat on the value of the wall resistance.

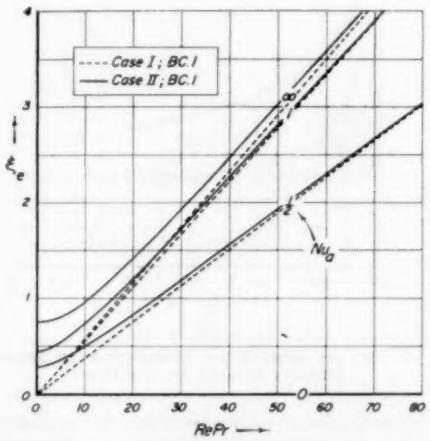


FIG. 10 EFFECT OF AXIAL FLUID CONDUCTION ON THERMAL-ENTRY LENGTH ξ_e FOR PARALLEL PLATES AS A FUNCTION OF PECLET NUMBER Pe = Re Pr

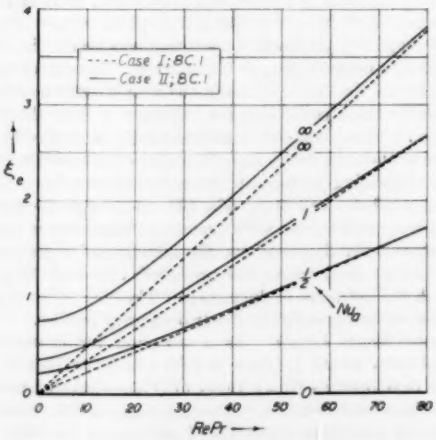


FIG. 11 EFFECT OF AXIAL FLUID CONDUCTION ON THERMAL-ENTRY LENGTH ξ_e FOR TUBES AS A FUNCTION OF PECLET NUMBER Pe = Re Pr

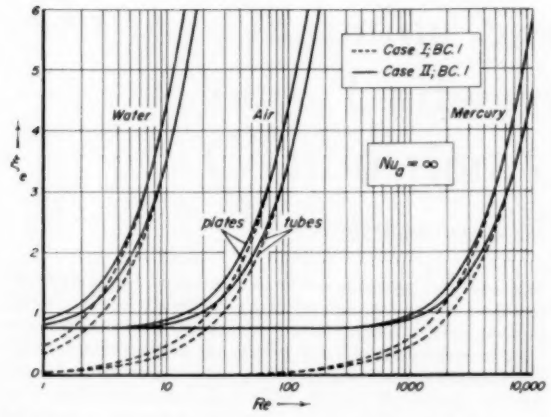


FIG. 12 THERMAL-ENTRY LENGTH ξ_e FOR THREE FLUIDS AS A FUNCTION OF REYNOLDS NUMBER RE

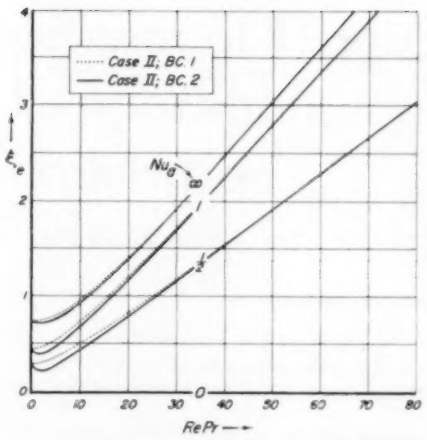


FIG. 13 DOWNSTREAM THERMAL-ENTRY LENGTH ξ_e IN PARALLEL PLATES WITH AXIAL FLUID CONDUCTION FOR BC.1 AND BC.2 AS A FUNCTION OF PECLET NUMBER PE = RE PR

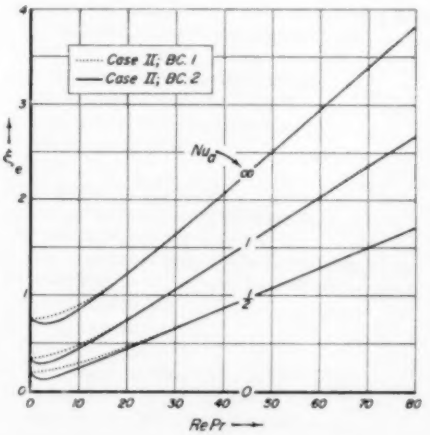


FIG. 14 DOWNSTREAM THERMAL-ENTRY LENGTH ξ_e IN TUBES WITH AXIAL FLUID CONDUCTION FOR BC.1 AND BC.2 AS A FUNCTION OF PECLET NUMBER PE = RE PR

An example of calculated downstream thermal-entry lengths for BC.2 is shown in Figs. 13 and 14 for parallel plates and tubes, respectively.[6] These results are for Case II (axial conduction included) and are compared here with the corresponding results for Case II, BC.1 from Figs. 10 and 11. For pure conduction (Pe = 0) the thermal-entry lengths are the same for both boundary conditions. However, in the case of BC.2 the minimum entry length does not occur, as it does for BC.1, at Pe = 0. That this minimum occurs at a rather low Peclet number is suggested in Fig. 2 wherein one notes that for Pe = 1 the downstream temperature is more established for BC.2 than for BC.1. Except for this small dip, the growth of the downstream thermal-entry lengths for the two boundary conditions is substantially the same at Peclet numbers of roughly 20 and above.

BIBLIOGRAPHY

1 "Heat Transfer," by M. Jakob, John Wiley & Sons, Inc., New York, N. Y., vol. 1, 1949.

2 "Heat Transfer in Laminar Flow Between Parallel Plates," by J. A. Prins, J. Mulder, and J. Schenk, *Applied Scientific Research*, section A, vol. 2, 1951, pp. 431–438.

3 "Heat Transfer in Laminar Flow Between Parallel Plates," by J. A. W. van der Does de Bye and J. Schenk, *Applied Scientific Research*, section A, vol. 3, 1953, pp. 308–316.

4 "Heat Transfer in Laminar Flow Through Cylindrical Tubes," by J. Schenk and J. M. Dumoré, *Applied Scientific Research*, section A, vol. 4, 1954, pp. 39–51.

5 "Heat Transfer in Laminar Flow Between Parallel Plates," by J. Schenk and H. L. Beckers, *Applied Scientific Research*, section A, vol. 4, 1954, pp. 405–413.

6 "Zur Theorie des Wärmeübergangs in laminarer Strömung," by M. V. Bodnarescu, *VDI-Forschungsheft* 450, Ed. B, vol. 21, 1955, pp. 19–27.

7 "Liquid-Metal Heat Transfer Coefficients," by R. N. Lyon, *Chemical Engineering Progress*, vol. 47, February, 1951, pp. 75–79.

8 "Heat Transfer in Laminar Flow in a Round Tube or Flat Conduit—The Graetz Problem Extended," by J. R. Sellars, M. Tribus, and J. S. Klein, Trans. ASME, vol. 78, 1956, pp. 441–448.

9 "Liquid-Metal Heat Transfer," by H. F. Poppendiek, Heat Transfer Symposium, University of Michigan Press, 1953, pp. 77–100.

10 "Laminar-Flow Forced Convection in Rectangular Tubes," by S. H. Clark and W. M. Kays, Trans. ASME, vol. 75, 1953, pp. 859–866.

11 "Numerical Solutions for Laminar-Flow Heat Transfer in Circular Tubes," by W. M. Kays, Trans. ASME, vol. 77, 1955, pp. 1265–1274.

[6] The downstream asymptotic Nusselt numbers are the same as for BC.1 in Fig. 8.

12 "Analysis of Laminar Forced-Convection Heat Transfer in Regions of Flat Rectangular Ducts," by E. M. Sparrow, NACA TN 3331, January, 1955.

13 "Heat Effects in Capillary Flow, I," by H. C. Brinkman, *Applied Scientific Research*, section A, vol. 2, 1951, pp. 120–124.

14 "Temperatur-, Zähigkeits-, und Reibungsverhältnisse in raschlaufenden Gleitlagern," by B. Dizioglu, Fifty Years Boundary Layer Research, I. R. Maxwell & Company, Ltd., London and New York, 1955, pp. 236–256.

15 "Analytical Investigation of Fully Developed Laminar Flow in Tubes With Heat Transfer With Fluid Properties Variable Along the Radius," by R. G. Diessler, NACA TN 2410, July, 1951.

16 "Conduction of Heat in Solids," by H. S. Carslaw and J. C. Jaeger, Oxford University Press, London, England, 1948.

17 "Conduction Heat Transfer," by P. J. Schneider, Addison-Wesley, Cambridge, Mass., 1955.

18 "On Convection of Heat," by H. A. Wilson, Proceedings of the Cambridge Philosophical Society, vol. 12, 1903–1904, pp. 406–423.

19 "Non-Uniform Heat Transfer to Fluids Flowing in Conduits," by V. J. Berry, Jr., *Applied Scientific Research*, section A, vol. 4, 1954, pp. 61–75.

On the Changing Size Spectrum of Particle Clouds Undergoing Evaporation, Combustion, or Acceleration[1]

By A. H. SHAPIRO[2] and A. J. ERICKSON,[3] CAMBRIDGE, MASS.

A theoretical treatment is given showing how the size distribution of a cloud of particles changes as the result of evaporation, combustion, or acceleration. The general differential equation governing the concentration of particles as a function of size, position, and time is formulated for one-dimensional duct-type flows. Solutions to the differential equation are then obtained for a number of special problems of interest to evaporation and combustion. When molecular transfers control, the equivalent mean diameter for evaporation or combustion of drops is found to be approximately constant with time. This suggests that the conventional model of a constant number of uniform drops of varying size be replaced by a new model having a varying number of uniform drops of constant size. The new model predicts a lower rate of evaporation or combustion than the conventional model.

Nomenclature

The following nomenclature is used in the paper:

A = cross-sectional area of duct
c_1, c_2, c_3 = constants of integration
c_p = specific heat at constant pressure of surrounding medium
$\mathfrak{D}$ = diameter of particle
$\mathfrak{D}^*$ = constant reference diameter equal to $\mathfrak{D}$ at the point $[\partial G/\partial \mathfrak{D}]_{t=0} = 0$
$\mathbf{D}$ = dimensionless diameter, $\mathfrak{D}/\mathfrak{D}^*$
D/Dt = signifies "substantial" differentiation with respect to time; i.e., while following a particle of fixed identity
E = a constant, $\mathfrak{L}\rho V_0/m_g c_p T_0$
$f_1(\mathfrak{D})$ = see Equation [13a]
$f_2(t)$ = see Equation [13a]
G = $dn/d\mathfrak{D}$, number of particles of size $\mathfrak{D}$ per unit volume of space and per unit interval of diameter
$\mathbf{G}$ = dimensionless form of G, $\mathbf{G} \equiv \mathfrak{D}^* G/n_0$
h = coefficient of heat transfer; coefficient of mass transfer
H = signifies an arbitrary function of the indicated argument
J = a function of diameter, $-\int d\mathfrak{D}/f_1(\mathfrak{D})$
K = constant in molecular growth rate law, $R = -K/\mathfrak{D}$
K_1 = constant in growth rate law for radiant heat transfer, $R = -K_1$
K_2 = constant in Stokes' law for terminal velocity, $u = K_2\mathfrak{D}^2$
$\mathfrak{L}$ = latent heat per unit mass of evaporating substance
m_g = mass of gas per unit volume of space
n = total number of particles per unit volume at time t
dn = number of particles per unit volume lying within the diameter interval from $\mathfrak{D}$ to $\mathfrak{D} + d\mathfrak{D}$
N = see Equation [25b]
q = rate of evaporation or combustion
R = growth rate of particle, $D\mathfrak{D}/Dt$
s = longitudinal distance along duct
S = see Equation [25b]
t = time
T = temperature of particle
T_g = temperature of surrounding medium
u = velocity of particle
V = total volume of particles in unit volume of space
x = mass fraction of particle cloud which has disappeared
α = see Equation [3]
β = see Equation [3]
γ = exponent in $R \sim \mathfrak{D}^{-\gamma}$
θ = dimensionless time variable, $Kt/\mathfrak{D}^{*2}$
θ_1 = dimensionless time variable, $K_1 t/\mathfrak{D}^*$
λ = thermal conductivity; mass diffusivity
ρ = mass density of particles
σ = transformed distance for observer moving with cloud, $\sigma = s - ut$
τ = temperature difference, $\tau \equiv T_g - T$
$(^-)$ = signifies quantities pertaining to appropriate mean size for same total mass and same rate of change of mass
$(\ )_u$ = signifies quantities for conventional model having fixed number of particles of uniform but varying size with correct initial mass and correct initial rate of change of mass
$(\ )_0$ = signifies quantities at $t = 0$

1 Introduction

The Problem. Many industrial processes require that a cloud of solid particles or liquid droplets interact with the gaseous (or, sometimes, liquid) phase in which they are dispersed. In certain of these processes, there is a spectrum of particle sizes, and, moreover, the particles change in size by reason of the interaction. Examples include (*a*) evaporation of a cloud of liquid droplets, (*b*) growth of a liquid cloud by condensation, (*c*) combustion of either solid-fuel particles or liquid-fuel droplets. Normally the rate of growth (taken of course in an algebraic sense) of each particle will depend on, among other things, the diameter of the particle itself.

[1] This work was partially sponsored by the Office of Naval Research, Department of the Navy. Reproduction in whole or in part is permitted for any purpose by the United States Government.

[2] Professor of Mechanical Engineering, Massachusetts Institute of Technology; presently Visiting Professor of Applied Thermodynamics, Engineering Laboratory, Cambridge University, Cambridge, England. Mem. ASME.

[3] Assistant Professor of Mechanical Engineering, Massachusetts Institute of Technology. Assoc. Mem. ASME.

Contributed by the Heat Transfer Division of The American Society of Mechanical Engineers and presented at the Heat Transfer and Fluid Mechanics Institute, Stanford University, Calif., June 21–23, 1956.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, May 11, 1956.

Under such conditions, the shape of the particle spectrum will change as time proceeds, and this naturally introduces considerable difficulty into the analysis of the problem. Even when there is no change in particle size, the spectral distribution of sizes may change; for example, if the cloud is accelerated, different sizes of particles will accelerate at different rates, and the differences in particle speeds will alter the relative concentrations per unit volume of the several particle sizes.

Object and Plan of Investigation. It is the purpose of this paper to develop some understanding of how the existence of a spectrum of particle sizes, coupled with a size-dependent growth rate, influences the processes described.

We shall investigate in what way a given initial spectrum changes shape. This requires formulation and solution of the governing differential equation. Then we shall look into the question of whether processes of the type considered can indeed be treated by means of a simple model comprising a cloud of particles of uniform "equivalent" size.

Most of the examples will refer to evaporation or combustion, with molecular transfer rates controlling. However, it is to be understood that the concepts and methods are equally applicable to other processes.

Previous Work. In virtually all theoretical investigations of evaporation, combustion, and so on, the actual cloud of particles is replaced by a simple model comprising a constant number of drops of uniform but changing size.

The actual size distribution in the combustion of pulverized coal was considered by Hottel and Stewart (1),[4] but their treatment of the problem, being essentially numerical, was rather cumbersome and not adapted to general use.

Probert (2) treated the size spectrum during combustion by accounting for current drops within a certain size interval as the remains of larger drops existing at the beginning of combustion. Although not specifically stated by Probert, one interesting result obtainable from the calculations was that the mean size of all droplets present in the combustion chamber in the steady state may, depending on circumstances, be either greater or smaller than the mean size of the injected droplets.

In the present paper, the governing equation of the particle spectrum is formulated in differential rather than integral form. Fortunately the equation is such that the general form of the solution for some cases may be found. The resulting analysis is therefore not only simple and straightforward, but applicable to a broad variety of practical problems.

2 Basic Concepts and Definitions

The Particle-Size Spectrum. Assuming that the actual histogram of the size distribution of particles may for purposes of analysis be replaced by a continuous curve, we define the particle concentration (or spectrum ordinate) G as $G \equiv dn/d\mathfrak{D}$, where dn is the number of particles per unit volume of space lying within the infinitesimal range of diameters between $\mathfrak{D}$ and $\mathfrak{D} + d\mathfrak{D}$. Then a graph of G versus $\mathfrak{D}$, Fig. 1(a), illustrates the size distribution in the cloud. The area under a narrow vertical strip, $Gd\mathfrak{D}$, represents dn, the number of particles per unit volume having diameters within the size range of the strip.

Dimensionless Representation. When the shape of the spectrum, rather than the actual values, is of interest, it is convenient to employ the dimensionless co-ordinates, $\mathfrak{D}^*G/n_0 \equiv \mathbf{G}$ and $\mathfrak{D}/\mathfrak{D}^* \equiv \mathbf{D}$, where $\mathfrak{D}^*$ is any convenient but fixed value of $\mathfrak{D}$ (for example, the value of $\mathfrak{D}$ corresponding to the initial value of G). These normalized co-ordinates are shown in brackets in Fig. 1(a).

Total Number of Particles per Unit Volume. To find the total number of particles per unit volume of space, we take the total area under the curve of Fig. 1(a), inasmuch as

$$n = \int dn = \int_0^\infty Gd\mathfrak{D} = \int_0^\infty \mathfrak{D}^*Gd(\mathfrak{D}/\mathfrak{D}^*) = n_0 \int_0^\infty \mathbf{G}d\mathbf{D} \quad [1]$$

Total Volume of Particles. Similarly, the total volume of particles per unit volume of space, is given by

$$V = \frac{\pi}{6}\int_0^\infty \mathfrak{D}^3 dn = \frac{\pi}{6}\int_0^\infty \mathfrak{D}^3 Gd\mathfrak{D} = \frac{\pi}{6}\mathfrak{D}^{*3}n_0 \int_0^\infty \mathbf{D}^3\mathbf{G}d\mathbf{D} \quad [2]$$

and is therefore found by integrating under the spectrum curve after the ordinates of the latter have been weighted by the factor $\mathbf{D}^3$.

The Growth Rate. The most important feature of the present analysis is the changing size of the particles. This is described by the growth rate

$$R \equiv D\mathfrak{D}/Dt$$

where the operator D/Dt signifies "substantial" differentiation, i.e., R is the rate of increase of particle diameter for a particle of fixed identity. For evaporation or combustion, R would have the negative values.

The value of R depends on the type of process (evaporation, combustion, etc.); on the physical properties of the particles and surrounding medium (velocity, density, viscosity, thermal conductivity, temperature difference, etc.); and, most important for our present purpose, on the diameter $\mathfrak{D}$ of the particular particle concerned.

Mean Particle Size. One may define any number of mean sizes (3) through the general formula

$$\bar{\mathfrak{D}}_{\alpha/\beta} \equiv \left[\frac{\int_0^\infty \mathfrak{D}^\alpha Gd\mathfrak{D}}{\int_0^\infty \mathfrak{D}^\beta Gd\mathfrak{D}}\right]^{\frac{1}{\alpha-\beta}} \quad [3]$$

where α and β are arbitrary numbers, and $\bar{\mathfrak{D}}_{\alpha/\beta}$ is the corresponding mean diameter. For example, with $\alpha = 3$ and $\beta = 2$, the value of $\bar{\mathfrak{D}}_{3/2}$ thus obtained is the so-called volume-surface mean diameter; i.e., the particle diameter whose ratio of volume to surface is the same as that of the actual particle cloud.

Equivalent Mean Size for Evaporation or Combustion. More pragmatically, we restrict the definition of mean size to one which is of some use. Our real aim is to treat the actual particle cloud as though it were composed of uniformly sized particles. For the processes under consideration, it is evident that the actual cloud and the model cloud must agree in two respects; they must, instantaneously, have (a) the same total mass and (b) the same rate of change of total mass. Using bars to denote quantities referring to the model cloud of uniform drops, these requirements may be expressed as

$$\int_0^\infty \mathfrak{D}^3 Gd\mathfrak{D} \equiv \bar{n}\bar{\mathfrak{D}}^3$$

$$\int_0^\infty \mathfrak{D}^2 RGd\mathfrak{D} \equiv \bar{n}\bar{\mathfrak{D}}^2\bar{R}$$

These may be solved simultaneously for $\bar{\mathfrak{D}}$ and n to give

$$\frac{\bar{\mathfrak{D}}}{\bar{R}} = \frac{\int_0^\infty \mathfrak{D}^3 Gd\mathfrak{D}}{\int_0^\infty \mathfrak{D}^2 GRd\mathfrak{D}} \quad [4]$$

[4] Numbers in parentheses refer to Bibliography at end of paper.

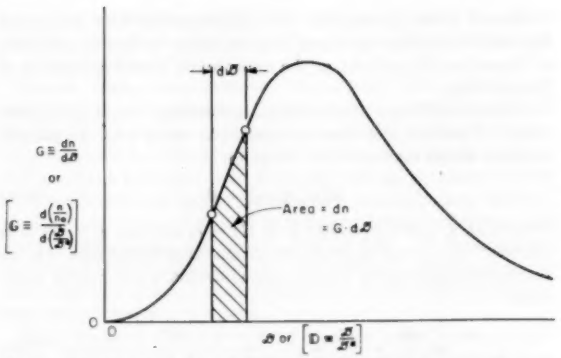


FIG. 1 (*a*) SPECTRUM OF PARTICLE SIZES, ILLUSTRATING NOMENCLATURE

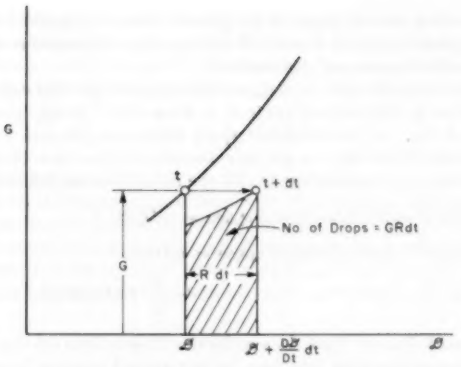


FIG. 1 (*b*) ILLUSTRATING GROWTH-FLUX FORMULA
(Shaded zone shows number of droplets which, during the time dt, have grown from sizes smaller than $\mathfrak{D}$ to sizes larger than $\mathfrak{D}$.)

$$\bar{n} = \frac{1}{\bar{R}^3} \frac{\left[\int_0^\infty \mathfrak{D}^3 R G d\mathfrak{D}\right]^3}{\left[\int_0^\infty \mathfrak{D}^3 G d\mathfrak{D}\right]^2} \quad \ldots\ldots\ldots\ldots [5]$$

At each instant, the model cloud containing $\bar{n}$ particles of uniform diameter $\bar{\mathfrak{D}}$ will have the same mass and the same rate of evaporation or combustion as the actual cloud. Here it is important to note, however, that $\bar{n}$ is not equal to the number of particles in the actual cloud; nor does it remain constant with the passage of time.

Examples of Practically Significant Mean Sizes. Often the dependence of R on $\mathfrak{D}$ may be approximated by $R \sim \mathfrak{D}^{-\gamma}$. The constant γ is zero when radiant heat transfer is controlling; it is equal to unity when molecular transfers control; and it is equal approximately to 0.2 when turbulent transfers control.

From Equation [4], the mean size is then

$$\bar{\mathfrak{D}}^{(1+\gamma)} = \frac{\int_0^\infty \mathfrak{D}^3 G d\mathfrak{D}}{\int_0^\infty \mathfrak{D}^{2-\gamma} G d\mathfrak{D}} \quad \ldots\ldots\ldots\ldots [4a]$$

The two limiting values of γ may be anticipated briefly as follows:

1 If R is independent of $\mathfrak{D}$, as would be the case when the process depends on radiant heat transfer, then $\gamma = 0$, and the appropriate mean is the volume-surface mean

$$\bar{\mathfrak{D}}_{\gamma=0} = \bar{\mathfrak{D}}_{1/2} = \frac{\int_0^\infty \mathfrak{D}^3 G d\mathfrak{D}}{\int_0^\infty \mathfrak{D}^2 G d\mathfrak{D}} \quad \ldots\ldots\ldots\ldots [4b]$$

2 If the process depends primarily on molecular transfers (i.e., the relative Reynolds number is very small), then $\gamma = 1$, reference (4), and the appropriate mean size is the 3/1 mean

$$\bar{\mathfrak{D}}^2{}_{\gamma=1} = \bar{\mathfrak{D}}_{3/1}{}^2 = \frac{\int_0^\infty \mathfrak{D}^3 G d\mathfrak{D}}{\int_0^\infty \mathfrak{D} G d\mathfrak{D}} \quad \ldots\ldots\ldots\ldots [4c]$$

3 DIFFERENTIAL EQUATION GOVERNING HISTORY OF PARTICLE SPECTRUM

Description of Model of Process. We consider here the flow of a discrete particle cloud in a duct or stream tube (the latter referring to the particle flow) of variable cross-sectional area A. The flow is considered one-dimensional to the extent that, over each cross-sectional area, the particle concentration G, the particle speed u, and the particle growth rate R are all uniform for each particle diameter $\mathfrak{D}$. G, u, and R are all considered to depend on particle size, longitudinal location, and time; e.g. $u = u(\mathfrak{D}, s, t)$.

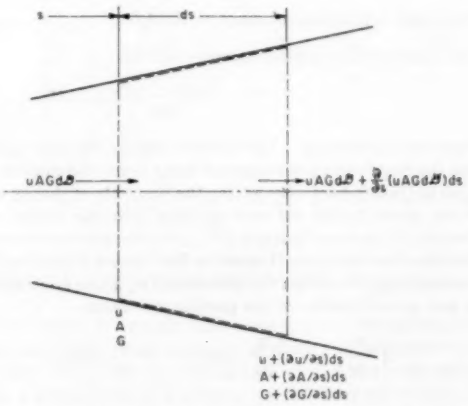


FIG. 2 CONTROL SPACE FOR FORMULATION OF DIFFERENTIAL EQUATION

The Growth Flux. A necessary preliminary to the differential equation is the derivation of an expression which indicates the flux rate at which, as a consequence of growth, particles cross from sizes smaller than $\mathfrak{D}$ to sizes larger than $\mathfrak{D}$.

Consider all the particles smaller than $\mathfrak{D}$ in a unit volume of space at time t. During the time interval dt, all the particles will have grown at their appropriate growth rates. The increase of diameter of those particles originally of size $\mathfrak{D}$ is $d\mathfrak{D} = Rdt$, and the number of particles grown larger than $\mathfrak{D}$ is therefore $GRdt$ (see Fig. 1*b*). Dividing this expression by dt, we obtain an expression for the growth flux

$$\left\{\begin{matrix}\text{Number of particles per unit time and} \\ \text{per unit volume becoming larger than } \mathfrak{D}\end{matrix}\right\} = GR \ldots [6]$$

Governing Differential Equation. We shall make a numerical accounting of only those droplets lying within the size range $\mathfrak{D}$ to $\mathfrak{D} + d\mathfrak{D}$ in the control space of Fig. 2. At the location s, let the area be A and let the properties corresponding to $\mathfrak{D}$ be denoted by u, G, and so on; at the exit location $s + ds$, each will at the same instant be larger by differential increments $(dA/ds)ds$, $(\partial u/\partial s)ds$, $(\partial G/\partial s)ds$, and so on.

The numerical accounting of droplets in the size range from $\mathfrak{D}$ to $\mathfrak{D} + d\mathfrak{D}$ requires a consideration of the convective fluxes into and

out of the control space, of the growth fluxes into and out of the size interval, and of the rate of change of particle number within the control space and size interval.

Convective Fluxes. At location s the number per unit volume of droplets in the size range $\mathfrak{D}$ to $\mathfrak{D} + d\mathfrak{D}$ is $Gd\mathfrak{D}$. Multiplying this by uA, i.e., by the volume of space swept out per unit time by droplets of this size, we get an expression for the convective flux entering, a corresponding one for the instantaneous flux leaving

$$\text{Rate of entry into control space} = uAGd\mathfrak{D} \quad [7a]$$

$$\text{Rate of exit from control space} = uAGd\mathfrak{D} + \frac{\partial}{\partial s}(uAGd\mathfrak{D})ds \quad [7b]$$

Growth Fluxes. Particles smaller than $\mathfrak{D}$ grow into the size range under consideration, and may, for our present purpose, be said to be created. Likewise, particles larger than $\mathfrak{D} + d\mathfrak{D}$ grow out of the size range and may be said to be destroyed. Taking note of Equation [6], we may accordingly write

$$\text{Birth rate within control space} = GRAds \quad [8a]$$

$$\text{Death rate within control space} = GRAds + \frac{\partial}{\partial \mathfrak{D}}(GRAds)d\mathfrak{D} \quad [8b]$$

Numerical Accounting. The instantaneous number of particles within the considered size range and lying inside the control space is equal to $(Gd\mathfrak{D})(Ads)$. By setting the time rate of change of this quantity equal to the net rate of entry into the control space (Equation [7a] minus Equation [7b]), plus the net rate of creation within the control space (Equation [8a] minus Equation [8b]), and simplifying, we obtain the differential equation governing the time and space histories of the particle spectrum

$$A\frac{\partial G}{\partial t} = -\frac{\partial}{\partial s}(uAG) - A\frac{\partial}{\partial \mathfrak{D}}(GR) \quad [9a]$$

or

$$\frac{1}{G}\left[\frac{\partial G}{\partial t} + u\frac{\partial G}{\partial s} + R\frac{\partial G}{\partial \mathfrak{D}}\right] = -\frac{\partial u}{\partial s} - \frac{u}{A}\frac{dA}{ds} - \frac{\partial R}{\partial \mathfrak{D}} \quad [9b]$$

Now, since $G = G(t, s, \mathfrak{D})$, we may write

$$\frac{dG}{dt} = \frac{\partial G}{\partial t}\frac{dt}{dt} + \frac{\partial G}{\partial s}\frac{ds}{dt} + \frac{\partial G}{\partial \mathfrak{D}}\frac{d\mathfrak{D}}{dt}$$

Specializing this for a group of particles of fixed identity, for which $ds/dt = u$, and $d\mathfrak{D}/dt = R$, we obtain an expression for the rate of change of concentration of particles of fixed identity, identified by the substantial derivative notation, in the form

$$\frac{DG}{Dt} = \frac{\partial G}{\partial t} + u\frac{\partial G}{\partial s} + R\frac{\partial G}{\partial \mathfrak{D}} = -G\left[\frac{\partial u}{\partial s} + \frac{u}{A}\frac{dA}{ds} + \frac{\partial R}{\partial \mathfrak{D}}\right] \quad [9c]$$

where the second equality is found by reference to Equation [9b].

Equation [9c] shows mathematically what might be deduced as well from purely physical considerations; i.e., that the spatial concentration of a certain group of particles of fixed identity (*a*) is decreased when the particles accelerate; (*b*) is decreased when the particles pass to a section of greater cross-sectional area; and (*c*) is decreased when the growth rate increases algebraically with diameter.

General Form of Solution. To determine the time and space histories of the size spectrum, it is necessary to find the solution of Equation [9], subject to the appropriate initial conditions of the problem.

Equation [9] is a linear partial differential equation of first order. The theory of such equations (5) shows that the general solution of this equation is of the form

$$H(v_1, v_2, v_3) = 0 \quad [10a]$$

where $v_1(t, s, \mathfrak{D}, G) = c_1$, $v_2(t, s, \mathfrak{D}, G) = c_2$ and $v_3(t, s, \mathfrak{D}, G) = c_3$ (in which c_1, c_2, and c_3, arbitrary constants of integration) are independent solutions of the associated ordinary differential equations

$$\frac{dt}{1} = \frac{ds}{u} = \frac{d\mathfrak{D}}{R} = \frac{-dG}{G\dfrac{\partial u}{\partial s} + \dfrac{uG}{A}\dfrac{dA}{ds} + G\dfrac{\partial R}{\partial \mathfrak{D}}} \quad [10b]$$

In practice, the solutions of Equations [10*b*] may be troublesome, and may indeed not be possible except in special cases where the variables may be separated readily.

We therefore proceed to the consideration of some special cases with the aim of discovering typical facts concerning the behavior-particle spectra.

4 The Uniformly Moving Cloud With Size Change

Specification of Model. Consider the case where all the particles move at the same constant speed in a duct of constant area. Such might occur when the particles were carried along in a fluid stream and all particles had settled down to the constant fluid velocity. The particle spectrum, however, changes as the result of evaporation, combustion, and so on. Then we have the simplifications that $dA/ds = 0$, $\partial u/\partial s = 0$, and $\partial u/d\mathfrak{D} = 0$.

Simplified Differential Equation. With the foregoing assumptions, Equation [9] now becomes

$$\frac{\partial G}{\partial t} + u\frac{\partial G}{\partial s} = -\frac{\partial}{\partial \mathfrak{D}}(GR) \quad [11]$$

Transformation to Observer Moving With Cloud. If we consider an observer moving with the particle cloud at the speed of the stream, the stream and the imbedded particle cloud will all appear stationary. Therefore, it is convenient to define a distance co-ordinate

$$\sigma \equiv s - ut$$

which is the relation between distance s in the fixed reference frame and distance σ in the moving observer's reference frame. Applying the usual procedure for interchange of variables from the t, s, $\mathfrak{D}$ system to the t, σ, $\mathfrak{D}$ system, we transform Equation [11] into

$$\left(\frac{\partial G}{\partial t}\right)_{\sigma,\mathfrak{D}} = -\left[\frac{\partial}{\partial \mathfrak{D}}(RG)\right]_{\sigma,t} \quad [12]$$

In fact, we may drop the subscript σ altogether if we remember that Equation [12] is valid only when the observer moves with the cloud and that the resulting equation applies to a group of particles of fixed identity.

Steady-Flow Case. It may be noted that Equation [12] includes the special case in which the process is steady in the stationary reference frame (i.e., $[\partial G/\partial t]_{s,\mathfrak{D}} = 0$), but of course nonsteady in the reference frame of the moving observer.

Motionless Cloud in Stationary Medium. Also covered by Equation [12] is another case of particular interest, that in which the particle cloud is motionless in a fixed space, the particles either growing or diminishing in size as time proceeds. The distance

co-ordinate σ is then evidently irrelevant, and may be struck from Equation [12]; this also may be seen more directly by setting $u = 0$ in Equation [11].

Growth Rate Considerations. Equations [10b] are indeed simplified, as we have seen, by dropping the terms $\partial u/\partial s$ and dA/ds. To obtain the solution even to simplified cases, however, we must know how the growth rate R depends on $\mathfrak{D}$ and t. Now R depends on t to the extent that the physical properties of the fluid medium (such as the temperature, concentration, thermal conductivity, and so on), and of the particles, partly govern the rates of evaporation, combustion, and the like, and these properties may all be functions of time. For example, the temperature difference forcing heat transfer depends on how much has already been evaporated or combusted.

Since the form of the $R - t$ relationship is not known in advance, it must be found by simultaneously solving the spectrum equations and the equations governing the changes in properties of the fluid medium and of the particles. Consequently, the solution must involve such auxiliary relations as the energy equation, stoichiometric equations, heat-transfer relationships, etc., inasmuch as these enter into the determination of the rates of interaction between the cloud and the fluid medium as well as of the properties of the fluid medium.

Case for Which $R = f_1(\mathfrak{D})f_2(t)$. Although the solution for G as a function of t and $\mathfrak{D}$ cannot be found for the general case until the $R - t$ relationship is known, we can avoid the simultaneous solution of the spectrum and fluid-medium equations by assuming that the growth rate can be written as

$$\frac{D\mathfrak{D}}{Dt} \equiv R = f_1(\mathfrak{D})f_2(t) \quad \ldots\ldots\ldots\ldots\ldots [13a]$$

wherein the function f_1 depends on $\mathfrak{D}$ alone and f_2 on t alone.

With this assumption we are now able to first solve the spectrum equations and then determine the complete relationships for the interaction between the particle cloud and the medium. The physical situation to which Equation [13a] applies is the case of a uniformly moving cloud and medium in which (i) all particle properties are the same except the size, and (ii) the properties of the fluid medium appear the same to each particle. Many real problems are approximately of this type; see, for instance, the example given in Section 6.

General Form of Solution for Uniformly Moving Cloud. Referring to Equation [10b] again, the solution to the dt, ds equation is now

$$s - ut = \sigma = c_1$$

the dt, $d\mathfrak{D}$ equation has the solution

$$J + \int_0^t f_2 dt = c_2$$

where

$$J(\mathfrak{D}) \equiv -\int d\mathfrak{D}/f_1 \quad \ldots\ldots\ldots\ldots\ldots [13b]$$

and the $d\mathfrak{D}$, dG equation integrates to

$$Gf_1 = c_3$$

Therefore, from Equation [10a], the complete solution to the problem must be of the form

$$G = \frac{1}{f_1} H\left[\sigma, \left(J + \int_0^t f_2 dt\right)\right] \quad \ldots\ldots\ldots [14]$$

where H is an arbitrary function of the arguments indicated.

Method of Determining Form of Arbitrary Function. To determine the precise form of the function H in a particular case, we must have f_1 and f_2 expressed algebraically in terms of $\mathfrak{D}$ and t, respectively, and the relationship between G, $\mathfrak{D}$, and σ must be known at some time, say $t = 0$. Putting this latter relationship into the form indicated by Equation [14], one obtains the solution at any other time merely through replacing J at $t = 0$ by

$$\left(J + \int_0^t f_2 dt\right)$$

at any other time.

Graphical Interpretation of Solution. The solution represented by Equation [14] has a simple graphical interpretation. Focusing attention on a particular part of the cloud, i.e., $\sigma = \text{const}$, we suppose that the spectral curve $G(\mathfrak{D})$ is known at time $t = 0$. Let this function be plotted (see dashed curve in Fig. 3) in the

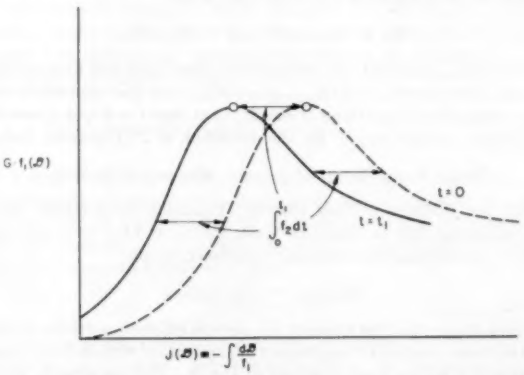


FIG. 3 GRAPHICAL INTERPRETATION OF SPECTRAL HISTORY FOR STATIONARY CLOUD WITH THE GROWTH-RATE LAW $R = f_1(\mathfrak{D})f_2(t)$

form Gf_1 versus J. At any later time t_1, the corresponding curve will be displaced without change of shape, as indicated by the solid curve of Fig. 3. According to Equation [14], the value of Gf_1 for a given value of J at time t_1 must be equal to the value of Gf_1 at $t = 0$ corresponding to a value of

$$J_{t=0} = J_{t_1} + \int_0^{t_1} f_2 dt$$

Hence the curve for time t_1 is simply displaced leftwards from that for $t = 0$ precisely by the magnitude

$$\int_0^{t_1} f_2 dt$$

Propagative Character of Solution. The foregoing features indicate a propagative behavior to the solution. For a given portion of the cloud, the product Gf_1 is seen to be constant for combinations of $\mathfrak{D}$ and t corresponding to a constant value of the function

$$J + \int_0^t f_2 dt$$

By separating variables and integrating between the limits $t = 0$, $\mathfrak{D} = \mathfrak{D}_{t=0}$, and $t = t$, $\mathfrak{D} = \mathfrak{D}$, we find from Equation [13a], for a particle of fixed identity

$$J_{\mathfrak{D}} + \int_0^t f_2 dt = J_{\mathfrak{D}(t=0)} \quad \ldots\ldots\ldots\ldots [13c]$$

where $J_{\mathfrak{D}}$ is the value of J at time t, and $J_{\mathfrak{D}(t=0)}$ is the value of J for the same particle at $t = 0$. In other words, particles of fixed identity have a $\mathfrak{D}$, t history specified by a constant value of

$$J + \int_0^t f_2 dt$$

which, by Equation [13c], is the value $J_{\mathfrak{D}(t=0)}$ associated with the particle at time zero. Then Equation [14] may be interpreted as meaning that the product Gf_1 remains constant for particles of fixed identity.

Thus the propagative behavior noted in the foregoing is associated with the fact that the initial concentration of particles in a certain size range controls the concentration of the same particles, but in successively different size ranges, at all later times.

Alternative Derivation of Solution by Physical Considerations. The propagative behavior also may be brought out by a derivation of Equation [14] based on more physical considerations. Consider a group of particles initially ($t = 0$) lying within a certain narrow size range, $d\mathfrak{D}_{t=0}$. Since the number of such particles at time t is equal to the number at $t = 0$, we may write

$$dn = G_{\mathfrak{D}(t=0)}d\mathfrak{D}_{t=0} = G_{\mathfrak{D}(t)}d\mathfrak{D}_t$$

where $d\mathfrak{D}_{t=0}$ and $d\mathfrak{D}_t$ are, respectively, the initial and final widths of the size interval, and $G_{\mathfrak{D}(t=0)}$ and $G_{\mathfrak{D}(t)}$ are the concentrations corresponding to particles of size $\mathfrak{D}_{t=0}$ at time $t = 0$ and of size $\mathfrak{D}_t$ at time t, respectively. By the definition of J (Equation [13b])

$$d\mathfrak{D}_{t=0} = -f_1(\mathfrak{D}_{t=0})dJ_{\mathfrak{D}(t=0)}; \quad d\mathfrak{D}_t = -f_1(\mathfrak{D}_t)dJ_{\mathfrak{D}(t)}$$

But, for particles of fixed identity, Equation [13c] holds. Differentiating this at constant time, we have $dJ_{\mathfrak{D}} = dJ_{\mathfrak{D}(t=0)}$. Now, assembling the foregoing equations, we get

$$(Gf_1)_{\mathfrak{D}(t)} = (Gf_1)_{\mathfrak{D}(t=0)}$$

which states that the product Gf_1, associated with particles of size $\mathfrak{D}$ at time t is equal to the product Gf_1 associated with the size $\mathfrak{D}_{t=0}$ belonging to the same particles at $t = 0$. But we already have seen that the identity of a certain group of particles is marked by a $\mathfrak{D}$, t relationship given by Equation [13c]. Consequently, we may conclude that the product Gf_1 is a function only of

$$J + \int_0^t f_2 dt$$

which is exactly what is claimed by Equation [14].

Case of Constant and Uniform Growth Rate. If the growth rate is the same, and constant, for all drops, the integral J of Equation [13] is simply $J = -\mathfrak{D}/R$, and the general solution to Equation [15] is $GR = H(\mathfrak{D} - Rt)$, which may be expressed alternatively as $G = H_1(\mathfrak{D} - Rt)$. This states exactly what the assumption of constant growth rate for all particles implies, namely, that the distribution curve of G versus $\mathfrak{D}$ marches across the graph with unchanged form, the displacement at any instant being Rt.

A possible application of this result is to problems in which the growth rate is controlled by radiant heat transfer. An energy balance in that case shows that R is independent of particle size.

Method of Treating Formation of New Drops. When the growth rate is positive, as in condensation, attention must be given to the fact that new drops may constantly be born. The type of solution already discussed applies only to those drops already present at time zero. Assuming that, at each instant, the birth rate of new drops per unit volume of space is known, we may equate this to the instantaneous value of GR corresponding to $\mathfrak{D} = 0$. This in turn yields $G_{\mathfrak{D}=0}$ as a function of time. Consequently, in respect to the new drops, the form of the arbitrary function in Equation [14] may be determined from a knowledge of its variation with t for the value of J corresponding to $\mathfrak{D} = 0$, and thus the general solution for the spectral history of the new drops may be found.

5 Evaporation or Combustion of a Cloud in a Very Large Medium

Description of Process. To illustrate more completely the method outlined, and also to obtain some practical results of value to an understanding of evaporation and combustion, we now take up the special case of a stationary cloud evaporating into, or burning in, a very large, stationary, gaseous medium. As indicated by the preceding section, the results also are applicable to each portion of fixed identity of a uniformly moving cloud. In this section we shall assume that the mass of the medium, relative to that of the cloud, is so large that the process does not alter the properties of the medium materially.

With this assumption, the growth rate for any portion of the cloud (i.e., for fixed σ) will depend only on diameter, and we may at once set $f_2(t) = 1$. Then, remembering that R is now only a function of $\mathfrak{D}$, Equation [14] becomes

$$G = \frac{1}{R} H(J + t) \ldots\ldots\ldots\ldots\ldots [15]$$

where H is an arbitrary function of the indicated argument. Alternatively, for any fixed portion of the cloud, Equation [15] states that the product GR (which varies only with $\mathfrak{D}$) depends only on the combined function of $\mathfrak{D}$ and t denoted by $(J + t)$. As before, the precise form of this dependency may be found by expressing $G(\mathfrak{D})$ at $t = 0$ in the form of Equation [15], and then replacing $J(\mathfrak{D})$ where it appears by $[J(\mathfrak{D}) + t]$.

The Molecular Growth-Rate Law. After the particles have reached a quasi-steady temperature, both evaporation and combustion are controlled by the processes of heat transfer and mass diffusion. If we assume no relative motion between the particles and the medium (or, more generally, that the Reynolds number based on the relative speed is small) dimensional considerations require that the coefficients of heat transfer and mass transfer follow the law $h\mathfrak{D}/\lambda = \text{const}$, where h is, respectively, the coefficient of heat transfer or mass transfer, and λ is, respectively, the thermal conductivity or mass diffusivity (4).

Since the rate of the process, and, consequently, the rate of disappearance of mass from each particle, is proportional to the coefficient h, to the surface area, and to differences in temperature and concentration, we may write that $\mathfrak{D}^2(D\mathfrak{D}/Dt) \sim h\mathfrak{D}^2$. Then noting that $h\mathfrak{D}$ is constant, we find that

$$R \equiv D\mathfrak{D}/Dt = -K/\mathfrak{D} = f_1 \ldots\ldots\ldots\ldots [16]$$

where K is a positive constant containing such quantities as pure numbers, the density and latent heat of the particles, the temperature difference, and the thermal conductivity and mass diffusivity of the medium. Thus we may evaluate

$$J = -\int \frac{d\mathfrak{D}}{f_1} = \int \frac{\mathfrak{D}d\mathfrak{D}}{K} = \mathfrak{D}^2/2K \ldots\ldots\ldots [17]$$

Moreover, integration of Equation [16] for a particle of fixed identity yields

$$\int_{\mathfrak{D}_{t=0}}^{\mathfrak{D}} \mathfrak{D}d\mathfrak{D} = -K\int_0^t dt; \quad \mathfrak{D}^2_{t=0} - \mathfrak{D}^2 = 2Kt \ldots [18]$$

where it is to be understood that $\mathfrak{D}_{t=0}$ is the diameter at $t = 0$ of a particle having the diameter $\mathfrak{D}$ at time t.

General Solution. According to Equation [15], the general solution to the problem considered here is

$$G = \mathfrak{D}H[\mathfrak{D}^2 + 2Kt] \ldots\ldots\ldots\ldots [19a]$$

Rearranging this in dimensionless form, we have

$$\mathbf{G} \equiv \mathfrak{D}^*G/n_0 = (\mathfrak{D}/\mathfrak{D}^*)H[(\mathfrak{D}/\mathfrak{D}^*)^2 + 2\theta] = \mathbf{D}H[\mathbf{D}^2 + 2\theta] \ldots\ldots [19b]$$

where $\theta \equiv Kt/\mathfrak{D}^{*2}$. The advantage of the form shown in Equation [19b] is that it removes the results from any particular scale

of time, size, or K, and expresses them in their most general form in terms of the dimensionless time, θ.

Graphical Method of Solution. A simple and rapid method of graphical solution is suggested by Equation [19b]. At time $t = 0$, the size distribution is plotted in the form of $\mathbf{G/D}$ versus $\mathbf{D}^2$. Then, for any other time $t = t_1$, the corresponding curve may be found merely by sliding the original curve leftwards by the amount $2\theta = 2Kt_1/\mathfrak{D}^{*2}$, as illustrated by Fig. 4. Then it is a

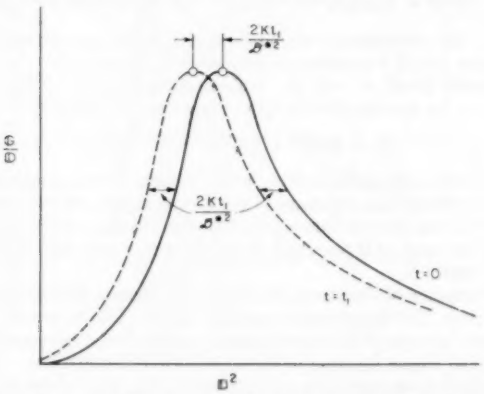


FIG. 4 GRAPHICAL SOLUTION FOR STATIONARY CLOUD EVAPORATING OR BURNING IN LARGE GASEOUS MEDIUM, WITH $R = -K/\mathfrak{D}$

simple matter to regraph the new curve for time t_1 in the conventional form of $\mathbf{G}$ versus $\mathbf{D}$ or G versus $\mathfrak{D}$.

Analytic Solution of Typical Case. These are several types of size distribution which may represent atomized sprays or milled solid particles (6). All have the features that the curve of G versus $\mathfrak{D}$ starts from zero, increases with $\mathfrak{D}$ and then, after reaching a maximum, approaches zero again asymptotically.

Initial Size Distribution. We consider now a specific type of distribution at $t = 0$ that embodies the foregoing features and that allows us to draw significant conclusions by simple calculations. The distribution postulated at $t = 0$ is represented by

$$\mathfrak{D}^*G = n_0(\mathfrak{D}/\mathfrak{D}^*)e^{-\frac{1}{2}(\mathfrak{D}/\mathfrak{D}^*)^2} \quad \text{or} \quad \mathbf{G} = \mathbf{D}e^{-\mathbf{D}^2/2} \quad [20a]$$

For very small values of $\mathfrak{D}$ this gives a linear distribution, while for large values the concentration approaches zero very rapidly. The constants in the equation are chosen so that the initially most populous particles (i.e., those of maximum G) are of size $\mathfrak{D}^*$, and the total number of particles initially is n_0. For the process considered, the instantaneous appropriate mean size is given by Equation [4c]. When the latter is evaluated for the distribution of Equation [20a], we get

$$\bar{\mathbf{D}}_{t=0} \equiv \bar{\mathfrak{D}}/\mathfrak{D}^* = \sqrt{3}$$

where it is understood that $\bar{\mathbf{D}}$ means, for this case, $\bar{\mathbf{D}}_{1/1}$.

Size Distribution at Any Time. Comparison of Equation [20a] with Equation [19b] shows that the form of H at $t = 0$ is

$$H_{t=0} = e^{-\mathbf{D}^2/2}$$

Consequently, at any other time t, the spectrum is given by

$$\mathbf{G} = \mathbf{D}e^{-\frac{1}{2}[\mathbf{D}^2+2\theta]} = e^{-\theta}\mathbf{D}e^{-\frac{1}{2}\mathbf{D}^2} \quad \ldots\ldots [20b]$$

This remarkable result shows that the size distribution considered remains unchanged in form, the concentrations of particles of all sizes being reduced by the constant factor $e^{-\theta}$.

Variation of Equivalent Mean Size. Performing the integrations of Equation [4c], we get the striking result that $\bar{\mathbf{D}} = \sqrt{3}$ at any time; that is, the appropriate mean size does not change, even though all particles are becoming smaller. The explanation of this seeming paradox is that the small particles grow smaller more rapidly than the large ones, and ultimately disappear entirely; as time proceeds, therefore, a relatively greater proportion of large particles remains.

Variation of Total Number of Particles. From Equations [1] and [20b], we find that the total number of particles decreases exponentially with time, i.e.

$$\frac{n}{n_0} = e^{-\theta} \quad \ldots\ldots [20c]$$

Variation of Fraction of Mass Disappeared. The fraction of the total volume of all particles lost, which is a measure of how much mass has disappeared through evaporation or combustion, is found from Equations [2] and [20b] as

$$x \equiv \text{fraction disappeared} = 1 - \frac{V}{V_{t=0}} = 1 - e^{-\theta} \quad [20d]$$

Variation of Rate of Evaporation or Combustion. The rate of evaporation or combustion is proportional to the rate of disappearance of total volume of all particles; i.e., $q \sim -dV/dt$. From Equations [1] and [20d], therefore, we find that this rate also varies exponentially with time

$$q/q_0 = e^{-\theta} \quad \ldots\ldots [20e]$$

Calculations for Conventional Model Containing Fixed Number of Particles of Uniform but Variable Size. In calculations of evaporation and combustion it is usual to assume a model in which there is a constant number of particles of uniform size, the initial uniform size being set equal to the appropriate mean size of the initial spectrum. To see how this model compares with the true state of affairs, we shall determine, for this model, the expressions corresponding to those of Equations [20d] and [20e] for the spectrum of Equation [20a]. Let $\mathfrak{D}_u$ and n_u denote, respectively, the uniform but variable particle size and the constant number of particles for this model. Then, by definition, $\mathfrak{D}_{u,\,t=0}/\mathfrak{D}^* = \sqrt{3}$. Furthermore, from Equation [18]

$$\left(\frac{\mathfrak{D}_u}{\mathfrak{D}^*}\right)^2 = \left(\frac{\mathfrak{D}_u}{\mathfrak{D}^*}\right)^2_{t=0} - 2\theta = 3 - 2\theta \quad \ldots\ldots [21a]$$

The fraction of the total volume of all particles which has disappeared at time t is given by

$$x_u = \left[1 - \frac{V}{V_{t=0}}\right]_u = 1 - \frac{(\mathfrak{D}_u/\mathfrak{D}^*)^3}{(\mathfrak{D}_u/\mathfrak{D}^*)_{t=0}^3} = 1 - \left(1 - \frac{2\theta}{3}\right)^{3/2} \quad \ldots\ldots [21b]$$

and the rate of evaporation or combustion (since $q \sim -dV/dt$) is, in proportion to the initial rate

$$(q/q_0)_u = \left(1 - \frac{2\theta}{3}\right)^{1/2} \quad \ldots\ldots [21c]$$

Comparison of Conventional Model With Actual Cloud. Equations [20d], [20e], [21b], and [21c] are graphed in Fig. 5. Since by definition, the initial mass and initial rate of loss of mass are identical for the spectrum and model, Fig. 5 shows directly the inaccuracies in the model. At the beginning of the process the two rates are equal and consequently the slopes of the x-curves are the same. As the process proceeds, however, the rate calculated for the model is at first greater than the true rate; this occurs because, in the true spectrum, the instantaneous mean size does not

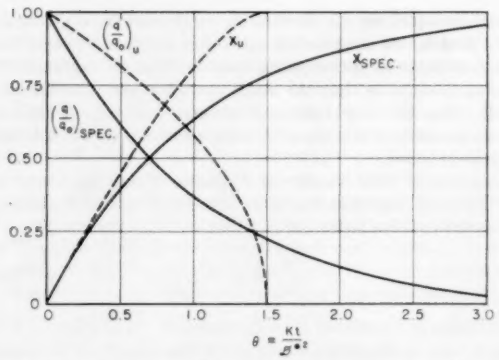


FIG. 5 COMPARISON OF FRACTION DISAPPEARED AND RATE OF DISAPPEARANCE, BOTH VERSUS TIME, FOR INITIAL SPECTRUM OF EQUATION [20a] (SOLID CURVES) AND FOR SIMPLE MODEL COMPRISING A CONSTANT NUMBER OF DROPS OF UNIFORM BUT VARIABLE SIZE (DASHED CURVES)

change, whereas in the simple model, all particles grow smaller and consequently acquire a larger growth rate (in the absolute sense). After about 75 per cent of the mass has disappeared (according to spectrum calculations) the rate calculated for the model becomes smaller than that for the spectrum; this happens because the excessively high rate at early times so reduces the particle diameter (and the surface area) of the model that ultimately a point must be reached where the rate becomes less than that of the spectrum. Notwithstanding the latter remark, at any instant the fraction disappeared is greater for the model than for the spectrum. Indeed, at $\theta = 1.5$ the particles of the model have entirely vanished, while in the spectrum they never entirely vanish. To summarize, the model of evaporation or combustion in which the spectrum is represented by a constant number of particles of uniform but variable size yields too high an estimate of how much mass will have disappeared at any given time. Table 1 gives some significant comparisons.

Proposed New Model of Particle Cloud. It is evident that the conventional model incorporates an unwarranted optimism concerning the time required for the process to occur. A more accurate model is suggested by the fact that the instantaneous $\bar{D}_{3/1}$ corresponding to Equation [20b] is constant. That is, the actual spectrum may be replaced by a cloud of particles of variable number but of uniform and constant size, with the number of such particles decreasing at a rate proportional to the rate of evaporation or combustion. Such a model, for the initial spectrum of Equation [20a] would give rates identical with those of the true spectrum. At first it may seem peculiar to imagine particles vanishing one by one, without diminution of diameter; but when it is recalled that the model is in any case a fiction to enable easy calculations, and that the number of particles in the model cloud has no physical association with the number in the true cloud, the seeming strangeness disappears.

It cannot be claimed that the model suggested in the foregoing is identically correct for other particle-size distributions, or for cases where R is not inversely proportional to $\mathcal{D}$. However, since the spectral curve of Equation [20a] is at least typical in shape, and since R generally varies with $\mathcal{D}$ to a power lying between 0 and —1, it seems safe to say that the conventional model incorporating a constant number of drops is generally overoptimistic in its estimate of rates, and that the new model proposed is often more realistic.

Results for Nukiyama-Tanasawa Distribution. To investigate the two points mentioned in the preceding paragraph, we now examine the Nukiyama-Tanasawa distribution, accepted as best representing that of liquid sprays atomized by air jets (6). The distribution equation at $t = 0$ is

$$\mathcal{D}^{*}G = 4n_0(\mathcal{D}/\mathcal{D}^{*})^2e^{-2(\mathcal{D}/\mathcal{D}^{*})} \quad \text{or} \quad \mathbf{G} = 4\mathbf{D}^2e^{-2\mathbf{D}} \quad [22a]$$

where the constants are so chosen that n_0 is the total number of droplets and $\mathbf{G}$ is a maximum when $\mathbf{D} = 1$, all at $t = 0$.

Results for $R = -K/\mathcal{D}$. Referring to Equation [19b], the solution for the size distribution at any time t is given by

$$\mathbf{G} = 4\mathbf{D}(\mathbf{D}^2 + 2\theta)^{1/2}e^{-2(\mathbf{D}^2+2\theta)^{1/2}} \quad [22b]$$

Inasmuch as the terms in θ cannot be brought out in a single factor, it follows that the shape of the spectrum changes with time. A few of the spectral curves are plotted in Fig. 6(*a*), which shows that the value of $\mathbf{D}$ for which $\mathbf{G}$ is a maximum increases slightly with time.

Evaluation of Equations [2] and [4c] for the size distribution of Equation [22] is somewhat lengthy, but the results can be expressed in terms of the second, fourth, and sixth derivatives of the modified Bessel function of the second kind of order zero. The results of these calculations are plotted in Fig. 6(*b*), which shows that the appropriate mean diameter for this distribution actually increases with time. This result is perhaps not too surprising after having seen the curves in Fig. 6(*a*). This increase in mean diameter, however, is not very large (about 17 per cent when the cloud is 75 per cent evaporated).

Fig. 6(*c*) shows the fraction of mass disappeared versus time for the spectrum as well as the corresponding curves for the conventional model with a constant number of drops of uniform but variable size, and for the new model of constant diameter but varying number. The curves illustrate clearly the error in using the conventional model for evaporating drops. The proposed new model does not give results that coincide exactly with the spectrum calculations; they are in error in the same direction as the conventional model. This is expected because $\bar{D}$ actually increases somewhat during evaporation and therefore the rate of change of mass is less for the spectrum. However, the fraction evaporated predicted by the proposed model is better than the conventional model both qualitatively and quantitatively.

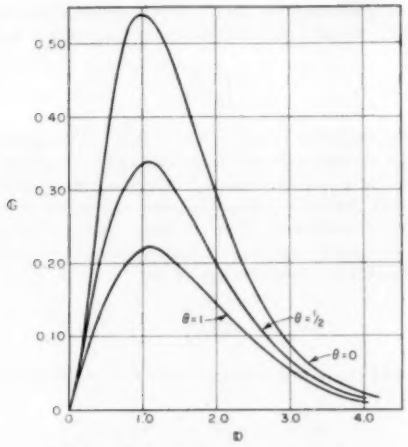
Results for R = Const. The results shown in Fig. 6 are for the case $R \sim \mathcal{D}^{-1}$. To see what effect the nature of the growth-rate law has on the results, Fig. 7 shows curves analogous to those of Fig. 6, for the same initial spectrum, but with $R = \text{const} = -K_1$. The two cases comprise the two extreme forms of the growth-rate law, for in the general relationship $R \sim \mathcal{D}^{-\gamma}$, γ lies between zero (radiant heat transfer controlling) and unity (molecular condition and diffusion controlling). Note that the appropriate mean diameter for $\gamma = 0$ is $\bar{D}_{3/2}$.

Fig. 7(*a*) shows the Nukiyama-Tanasawa spectrum as a whole moving horizontally towards the origin. This result was previously derived for any spectrum if the growth rate is constant and uniform. In Fig. 7(*b*), the mean diameter $\bar{D}_{3/2}$ is shown as a function of x. In this case the mean diameter decreases as the drops evaporate. Finally, in comparing the fraction evaporated curves versus θ_1 in Fig. 7(*c*), we see that once more the conventional model is overoptimistic about evaporation rates. The proposed model is also in error, but the magnitude of the error is smaller, and in the opposite direction. The proposed model again matches the spectrum result qualitatively over the entire range.

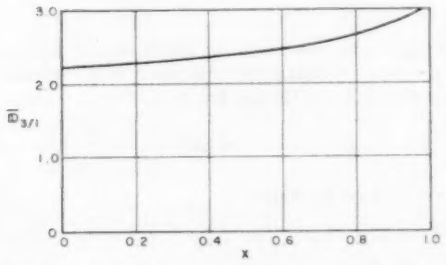
Spectrum With Finite-Diameter Range. In the preceding examples the diameter range for each spectrum ranged from zero to

TABLE 1 COMPARISON OF MODEL AND SPECTRUM

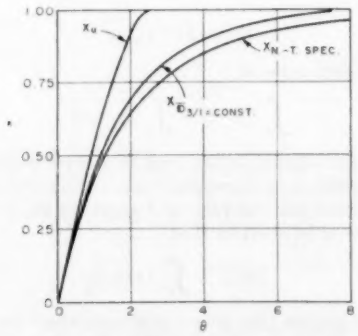
	Model	Spectrum
Per cent disappearance for $\theta = 0.555$	50	42.6
Per cent disappearance for $\theta = 0.905$	75	59.5
θ for 50 per cent disappeared	0.555	0.694
θ for 75 per cent disappeared	0.905	1.386
θ for 100 per cent disappeared	1.5	∞



(a) Spectral distributions at various times



(b) Variation of equivalent mean diameter with fraction evaporated or burned



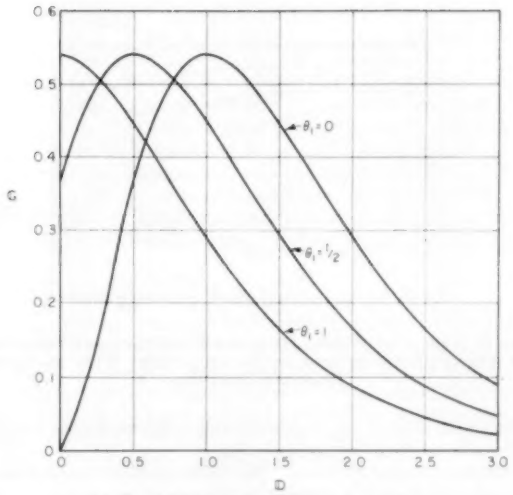
(c) Fraction evaporated or burned versus time for the spectrum, the conventional model, and the new model

FIG. 6 RESULTS FOR NUKIYAMA-TANASAWA SPECTRUM OF EQUATION [22a], WITH $R = -K/\mathfrak{D}$

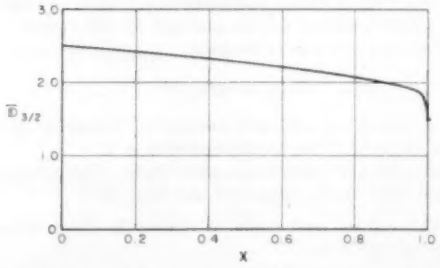
infinity. As a consequence, an infinite amount of time is required to evaporate all the droplets. Since the upper diameter limit is finite in any actual case, we now investigate this effect using the molecular growth-rate law of Equation [16] and the initial droplet spectrum of Equation [20a]. Thus at $t = 0$ we have

$$G = \begin{cases} \mathbf{D}e^{-\mathbf{D}^2/2} & 0 \le \mathbf{D} \le \mathbf{D}_{max(t=0)} \\ 0 & \mathbf{D} > \mathbf{D}_{max(t=0)} \end{cases}$$

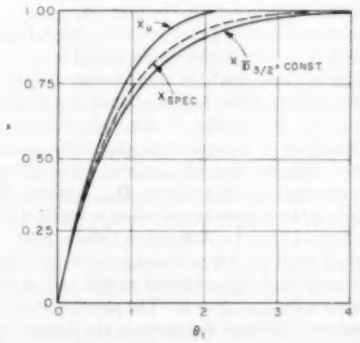
The correct mean diameter is given by



(a) Spectral distributions at various times



(b) Variation of equivalent mean diameter with fraction evaporated or burned



(c) Fraction evaporated or burned versus time for the spectrum, the conventional model, and the new model

FIG. 7 RESULTS FOR NUKIYAMA-TANASAWA SPECTRUM OF EQUATION [22a], WITH CONSTANT VALUE OF R

$$\bar{D}^3 = \frac{\int_0^{D_{max}} D^4 e^{-D^2/2} dD}{\int_0^{D_{max}} D^2 e^{-D^2/2} dD}$$

where D_{max} is not constant but can be expressed as

$$D_{max}^2 = D_{max(t=0)}^2 - 2\theta$$

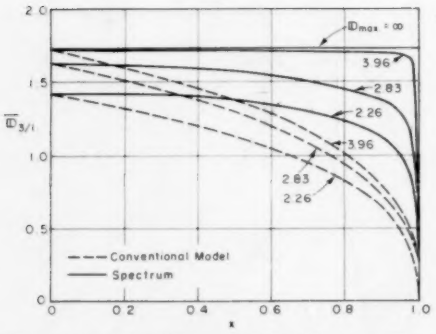


FIG. 8 VARIATION OF MEAN DIAMETER WITH FRACTION EVAPORATED OR BURNED FOR SPECTRUM OF EQUATION [20a] WITH AN UPPER DIAMETER LIMIT $\mathbf{D}_{max}$

The results of these calculations as well as the results of using the conventional model are shown in Fig. 8.

Choice of Initial Spectrum. The initial particle distribution G for any case is usually obtained by fitting a curve to experimental data. This procedure suggests trying to fit a curve of certain algebraic form such that later calculations for $\overline{\mathfrak{D}}$ yield the result $\overline{\mathfrak{D}}$ = const. This objective can be achieved for the case of $f_1(\mathfrak{D}) \sim -\mathfrak{D}^{-\gamma}$ if we fit the data to the form

$$G_{t=0} = a\mathfrak{D}^{\gamma}e^{-b\mathfrak{D}^{\gamma+1}}$$

where a and b are arbitrary constants. Equation [20a] is a specific example of this for the case of $\gamma = 1$.

Conclusions of Droplet-Evaporation Study. The general results obtained from study of Figs. 6, 7, and 8 are that:

1 The conventional model is optimistic about evaporation rates of the spectrum.

2 The proposed model is more accurate in predicting evaporation rates, at least for $\mathbf{D}e^{-\mathbf{D}^2}$ and $\mathbf{D}^2e^{-\mathbf{D}}$ spectra, for the range of growth rates from radiation controlling ($\gamma = 0$) to molecular phenomena controlling ($\gamma = 1$) the process.

3 If only a small part of the mass of the droplet spectrum is to be evaporated, it is immaterial which model is used because the diameter variation is so small ($\mathfrak{D} \sim$ cube root of mass remaining).

4 For evaporation of the last remnants of a droplet spectrum with $\mathbf{D}_{max} = \infty$, the results of the conventional model differ radically from those of the spectrum and of the proposed model. The latter two indicate that 100 per cent evaporation is reached only as t approaches ∞. If, however, $\mathbf{D}_{max}$ is finite, the spectrum analysis shows 100 per cent evaporation is reached in finite time while the proposed model still requires infinite time. Except for this question of time for 100 per cent evaporation, the difference between the results of the proposed model and of the spectrum analysis is small if $\mathbf{D}_{max(t=0)} > 3$. The physical reason for this behavior is that very little of the mass in the droplet spectrum we are considering is in the range from $\mathbf{D} = 3$ to $\mathbf{D} = \infty$. If the question of time for 100 per cent evaporation is of importance, the answer can easily be obtained by considering just the evaporation of the largest drop.

6 EVAPORATION OR COMBUSTION OF A STATIONARY CLOUD IN A "SMALL" MEDIUM

Form of Growth-Rate Law. When the growth rate depends on time, as for example when evaporation or combustion modifies the temperature and other properties of the medium, it is not valid to let $f_2(t) = 1$ in Equation [13a].

In this event, if we retain the growth-rate law $R = f_1(\mathfrak{D})f_2(t)$, with $f_1(\mathfrak{D}) \sim \mathfrak{D}^{-1}$, Equation [14] shows that all the results presented in Equations [19], [20], and [22] and in Figs. 5 and 6 remain valid, provided only that wherever t appears it must be replaced by

$$\int_0^t f_2(t)dt$$

Case of Evaporation in a Gaseous Medium. The details of completing the solution will vary from case to case, and may best be illustrated by a specific example. Consider evaporation into a cloud of gas, with the gas temperature falling by virtue of the latent-heat absorption. Then an energy balance for a droplet (neglecting sensible-heat effects), equating the rate of heat transfer to the rate of latent-heat change, yields

$$h\pi\mathfrak{D}^2\tau = -\rho\mathfrak{L}\frac{\pi}{2}\mathfrak{D}^2(D\mathfrak{D}/Dt)$$

whence, with the postulate of molecular heat conduction ($h\mathfrak{D}/\lambda = 2$) we get

$$R = D\mathfrak{D}/Dt = -4\lambda\tau/\rho\mathfrak{L}\mathfrak{D} \quad \ldots\ldots [23a]$$

Neglecting the variation of λ with time, we may express this as

$$R = f_1(\mathfrak{D})f_2(t) \text{ where } f_1 \equiv -4\lambda\tau_0/\rho\mathfrak{L}\mathfrak{D} \text{ and } f_2 \equiv \tau/\tau_0$$

An energy balance between the entire droplet cloud and the entire gas mass, equating the total latent heat supplied to the decrease of enthalpy of the gas, gives

$$m_g c_P(T_{g0} - T_g) = \rho\mathfrak{L}V_0\left(1 - \frac{V}{V_0}\right) \quad \ldots\ldots [23b]$$

Moreover we have by definition

$$\tau \equiv T_g - T = (T_{g0} - T) - (T_{g0} - T_g) = \tau_0 - (T_{g0} - T_g)$$

and thus we obtain

$$\tau = \tau_0[1 - Ex] \quad \ldots\ldots [23c]$$

where

$$E \equiv \rho\mathfrak{L}V_0/m_g c_P\tau_0$$

Now, for each value of

$$\int_0^t f_2(t)dt = \int_0^t (\tau/\tau_0)dt$$

the value of x is known either from an analytical solution (e.g., Equation [20d]) or, if necessary, from a numerical integration, keeping in mind that the value of θ appearing in, say, Equation [20d], is now to be interpreted as

$$(K/\mathfrak{D}^{*2})\int_0^t (\tau/\tau_0)dt$$

Moreover, Equation [23c] gives a relationship between τ/τ_0 and x, assuming of course that the constants τ_0, $\mathfrak{L}$, ρ, c_P, and V_0/m_g are all known to begin with. Consequently, we have in hand a relationship between τ/τ_0 and $\int(\tau/\tau_0)dt$. Then, from the expression

$$\tau = \frac{d}{dt}\int_0^t \tau dt = \frac{d\tau}{dt}\frac{d}{d\tau}\int_0^t \tau dt$$

we may separate variables and integrate between $t = 0$, $\tau = \tau_0$, and $t = t$, $\tau = \tau$, thus getting the identity

$$t \equiv \int_1^{\tau/\tau_0} \frac{\dfrac{d}{d(\tau/\tau_0)}\displaystyle\int_0^t (\tau/\tau_0)dt}{(\tau/\tau_0)} d(\tau/\tau_0) \quad \ldots\ldots [23d]$$

Since the differentiation and integration required by the right-hand side of this equation may be performed analytically or

graphically with the help of the aforementioned relationship between

$$\tau/\tau_0$$

and

$$\int_0^t (\tau/\tau_0)dt$$

we may solve for t as a function of τ/τ_0, either analytically or graphically. Then, since x is connected with τ/τ_0 through Equation [23c], we have our solution finally in the desired form of x versus t.

Specific Example. For example, with the rather simple algebraic relationship of Equation [20d], all the operations may be performed analytically and yield finally, for the relationship between fraction evaporated and time

$$t = \frac{\rho \mathfrak{L} \mathfrak{D}^{*2}}{4\lambda\tau_0(E-1)} \ln \frac{1-x}{1-Ex} \quad \ldots\ldots\ldots\ldots [24]$$

With other spectra the operations just indicated may have to be carried out numerically or graphically.

7 Steady-State Evaporation or Combustion in Ducts

Specification of Model. Let us now consider the case of evaporation or combustion in steady flow in a duct such that, at every station, all properties are constant, i.e., $\partial/\partial t = 0$.

Furthermore, let us suppose that the particles are very small and therefore that all have virtually the same speed as the medium at all points, i.e., $u = u(s)$ only. For the same reasons as given in the preceding section, let the growth rate be represented as $R = f_1(\mathfrak{D})f_2(s)$.

General Solution. With the foregoing assumptions, the application of Equation [10] yields, for the general solution to the problem

$$G = e^{-N}H(J+S) \quad \ldots\ldots\ldots\ldots\ldots [25a]$$

where

$$J \equiv \int (1/f_1)d\mathfrak{D}; \quad S \equiv \int (f_2/u)ds;$$

$$N \equiv \int \frac{\dfrac{\partial u}{\partial s} + \dfrac{u}{A}\dfrac{dA}{ds} + f_2 f_1'}{f_1 f_2} d\mathfrak{D} \quad \ldots\ldots [25b]$$

and H is an arbitrary function of the argument $J + S$. To evaluate the N integral, it is necessary to express s as a function of $\mathfrak{D}$ through the equation $J + S = J_{s=0}$, where $J_{s=0}$ is a constant of integration equal to the value of J at $s = 0$.

The actual treatment for particular cases is similar to that of the preceding section, where $\mathfrak{D}$ and t were the independent variables, except that now $\mathfrak{D}$ and s are the variables.

Special Case for Which uA = Const. When the density of the medium is not appreciably altered by the evaporation or combustion, and when the particles follow closely the motion of the medium, continuity requires that the product uA be nearly constant. Assuming that it is in fact constant, the integral N becomes

$$N = \ln f_1$$

and the solution takes the form

$$Gf_1 = H(J+S) \quad \ldots\ldots\ldots\ldots\ldots [26]$$

which bears a striking similarity to Equation [14].

General Solution for Mass Fraction Evaporated. The general solution for the spectral distribution as a function of distance, given by Equation [25], may be very difficult to carry through. However, should it be of interest only to discover how the fraction evaporated or burned varies with time, simple physical reasoning shows that the results of the preceding section (embodied, for example, in Equation [20d] and Fig. 6) are immediately applicable. If, as in the case with the assumptions employed here, each portion of the particle cloud is associated with a given mass of the medium, an observer moving with a given portion of cloud and medium would see the same events that would be seen if the cloud and medium were stationary. The only formula needed is that relating the distance moved by the observer to the time elapsed, and this is easily furnished by $s = \int u dt$. With this transformation, all the results of the preceding section connecting fraction disappeared with time may be used in the present section for connecting fraction disappeared with distance.

8 Steady Flow, With No Change in Size

Model of Process. Suppose that a cloud of particles travels in steady flow down a duct, with no interactions which might cause any particle to change in size. Then the concentration G might change because of (a) a change in cross-sectional area of the duct or (b) a change in the speed u of the particles of size $\mathfrak{D}$ in question.

General Solution. With the foregoing assumptions, $\partial/\partial t = 0$ and $R = 0$. Accordingly Equations [10b] are

$$\frac{ds}{u} = \frac{d\mathfrak{D}}{0} = \frac{-dG}{(G/A)\dfrac{\partial}{\partial s}(uA)}$$

The integrals of these equations are

$$\mathfrak{D} = c_1$$

and

$$GAu = c_2$$

and therefore the solution is

$$GAu = f(\mathfrak{D})$$

where $f(\mathfrak{D})$ signifies some function of $\mathfrak{D}$.

Physical Meaning of Solution. This solution has a simple physical meaning. Since G is the number concentration per unit volume, GAu is the number of particles in the size range from $\mathfrak{D}$ to $\mathfrak{D} + d\mathfrak{D}$, crossing each section per unit time. Consequently, the solution states that, for any particular size group, the number flux is the same at all cross sections. It would have been possible, therefore, to write down the foregoing solution simply by noting that the number of particles in each size range is conserved.

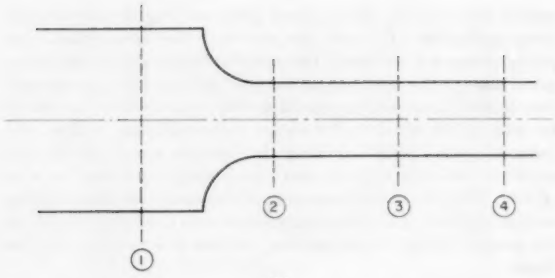
We note from the solution that the number concentration per unit volume G tends to decrease if either A or u increases.

Example of Motion in Incompressible Medium, With Slow Changes in Cross Section. Suppose that the particles are traveling in an incompressible medium, and that the changes of cross section are so moderate that all particles follow exactly the motion of the medium. The product uA is constant and it follows that the concentration curve $G(\mathfrak{D})$ is the same at all cross sections.

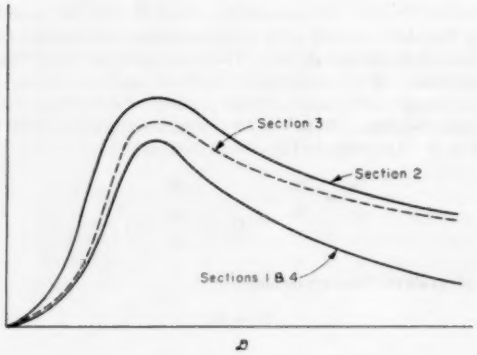
Example of Motion in Incompressible Medium, With Rapid Contraction of Cross Section. By way of contrast, let the medium, assumed as incompressible, accelerate through an abrupt decrease in cross section, as in Fig. 9(a). For reasons of particle inertia the smaller particles will accelerate from 1 to 2 more rapidly than the larger ones. Ultimately, however, all sizes of particles will arrive at the same speed at section 4. Furthermore, the reasoning given previously requires the spectrum at 4 to be the same as at 1. These arguments show that the spectral curves for sections 1, 2, 3, and 4 will change in the general way indicated by Fig. 9(b).

9 Free-Falling and Evaporating Droplet Cloud

Model and Solution of Process. Imagine a cloud of falling and evaporating drops. Assuming a steady state, we may set $\partial/\partial t = 0$.



(a)



(b)

FIG. 9 CHANGE OF SPECTRAL DISTRIBUTION WHEN PARTICLES CARRIED BY INCOMPRESSIBLE FLUID PASS THROUGH CONTRACTION IN DUCT

Moreover, if each drop is very nearly at its own terminal velocity at all times we may suppose u to be a function of $\mathfrak{D}$ only, and thus $\partial u/\partial s = 0$. If, as well, the surrounding medium is very large, we may suppose R to depend on $\mathfrak{D}$ alone. Then, noting that $dA/ds = 0$, Equations [10b] become

$$\frac{ds}{u} = \frac{d\mathfrak{D}}{R} = -\frac{dG}{G(dR/d\mathfrak{D})}$$

The integral of the ds, $d\mathfrak{D}$ equation is

$$s - S = c_1 \quad \text{where} \quad S(\mathfrak{D}) = \int \frac{u(\mathfrak{D})}{R(\mathfrak{D})}\, d\mathfrak{D}$$

and the integral of the $d\mathfrak{D}$, dG equation is

$$GR = c_2$$

Consequently the general solution is

$$GR = H(s - S)$$

in which the arbitrary function H is found from a knowledge of the function GR at the location $s = 0$.

Example for Very Small Droplets (Low Reynolds Number). To illustrate, suppose the droplets are so small that the evaporation occurs by molecular diffusion and the terminal velocity is set by Stokes' law of drag. These lead to the laws

$$R = -K/\mathfrak{D}; \quad u = K_2\mathfrak{D}^2$$

where K and K_2 are dimensional constants depending on the properties of the droplets and the medium. The function S may now be evaluated as

$$S \equiv \int \frac{K_2\mathfrak{D}^2}{-K/\mathfrak{D}}\, d\mathfrak{D} = -\frac{K_2}{4K}\,\mathfrak{D}^4$$

and the general solution showing the spectral distribution as a function of s is seen to be

$$G = \mathfrak{D}H\left(s + \frac{K_2}{4K}\,\mathfrak{D}^4\right)$$

Therefore, with a knowledge of how $G/\mathfrak{D}$ depends on $\mathfrak{D}^4$ at $s = 0$, it is a simple matter to find $G(\mathfrak{D})$ at any other value of s.

BIBLIOGRAPHY

1 "Space Requirement for the Combustion of Pulverized Coal," by H. C. Hottel and I. McC. Stewart, *Industrial and Engineering Chemistry*, vol. 32, 1940, pp. 719–730.

2 "The Influence of Particle Size and Distribution in the Combustion of Oil Droplets," by R. P. Probert, *Philosophical Magazine*, vol. 37, 1946, pp. 94–105.

3 "Droplet Size Distribution in Sprays," by R. A. Mugele and H. D. Evans, *Industrial and Engineering Chemistry*, vol. 43, 1951, pp. 1317–1324.

4 "Evaporation from Drops," by W. E. Ranz and W. R. Marshall, Jr., *Chemical Engineering Progress*, vol. 48, 1952, pp. 141–146; vol. 48, 1952, pp. 173–180.

5 "Advanced Calculus for Engineers," by F. B. Hildebrand, Prentice-Hall, Inc., New York, N. Y., 1949, pp. 368–378.

6 "Atomization of Liquids in High Velocity Gas Streams," by H. C. Lewis, D. G. Edwards, M. J. Goglia, R. I. Rice, and L. W. Smith, *Industrial and Engineering Chemistry*, vol. 40, 1948, pp. 67–74

Discussion

MYRON TRIBUS.[5] This paper represents a welcome addition to the growing literature on sprays. The evidence that the approximations suggested in the paper compare favorably with the calculation results from more expensive computational methods is reassuring.

The writer would like to illustrate a particular application which has been made of one of the many ideas in this paper. It is shown that,[6] for practical purposes, shortly after injection of a fine spray of a pure liquid into a gaseous atmosphere, each drop comes to the wet-bulb temperature of the system. Using the symbols of the paper (except letting $S = r^2$), it follows that all drops will follow the growth law

[5] Associate Professor of Engineering, University of California, Los Angeles, Calif. Mem. ASME.

[6] "A Method for Calculating the Rate of Evaporation and the Change in Drop Size Distribution for Pure Sprays Injected into Unsaturated Air," by M. Tribus, J. Klein, and J. Rembowski, Project M992-C, Engineering Research Institute, Ann Arbor, Mich., May, 1952 (available via microfilm or photostat).

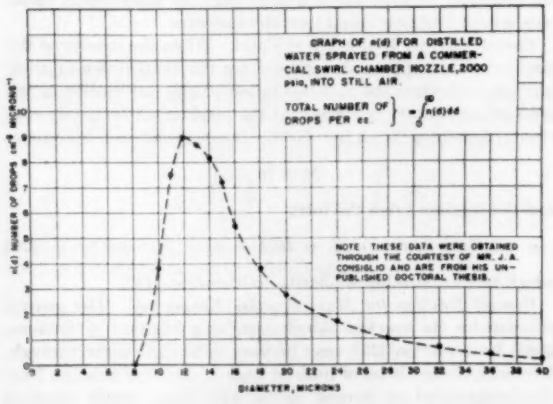


FIG. 10

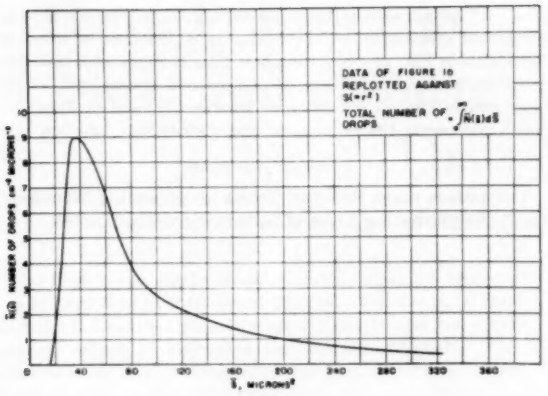


FIG. 11

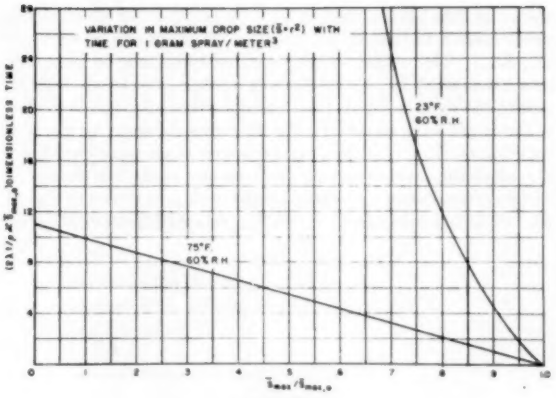


FIG. 14

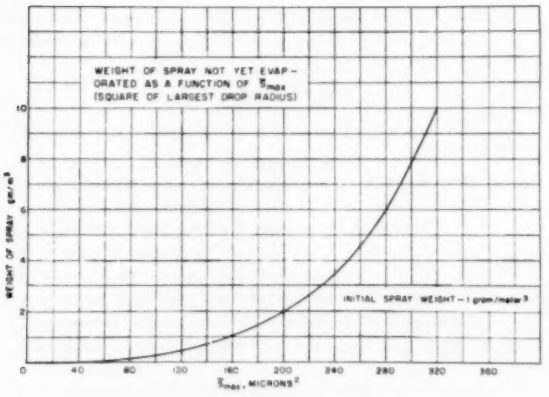


FIG. 12

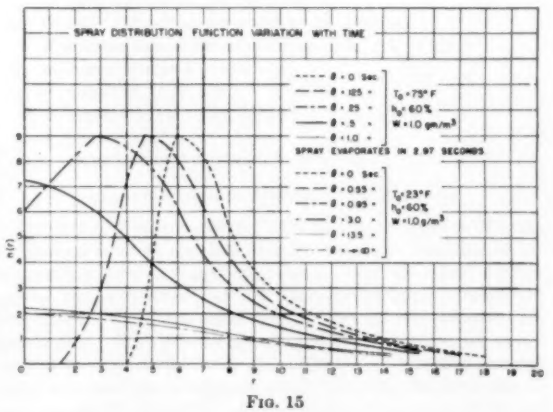


FIG. 15

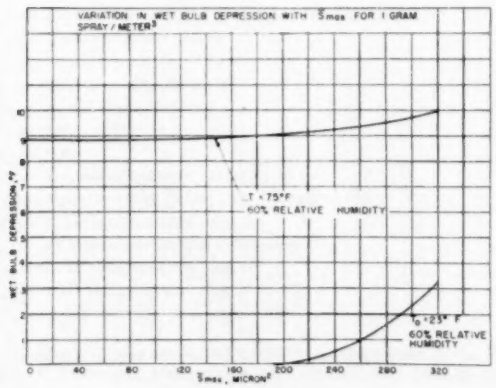


FIG. 13

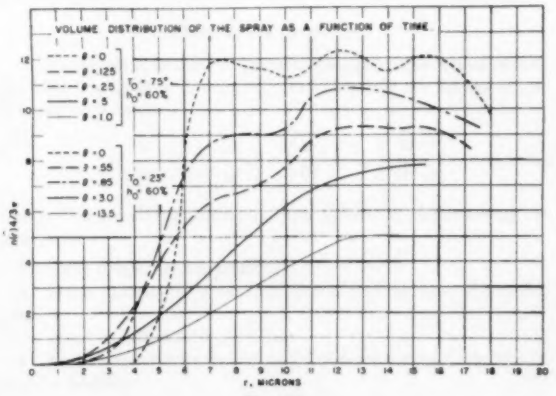


FIG. 16

$$\frac{dS}{dt} = \frac{dr^2}{dt} = \frac{2\lambda}{\rho\mathfrak{L}}(T_{wb} - T_{db}) \quad \ldots\ldots\ldots\ldots [27]$$

Instead of the distribution function $G(r)$ utilized in Equation [1] of the paper, consider a function $\bar{G}(\bar{S})$ such that the number of drops having values of r between $\bar{S}$ and $\bar{S} + d\bar{S}$ is $\bar{G}(\bar{S})d\bar{S}$. In view of Equation [27] herewith, it follows that $\bar{G}(\bar{S})$ does not change shape with time, since $d\bar{S}/dt$ is the same for all droplets.

The total weight of spray is given by

$$\int_0^\infty \rho(4/3)\pi\bar{S}^{3/2}\bar{G}(\bar{S})d\bar{S}$$

In the case of the evaporation from a water spray injected into unsaturated air, the variations in spray character are shown in Figs. 10 through 16, taken from reference.[6] Fig. 10 shows a typical drop-size spectrum plotted according to $G(r)$. Fig. 11 shows

the same data plotted with $\bar{G}(\bar{S})$. In practice the integration is never carried to infinity but stops at the maximum value of $\bar{S}$; i.e., $\bar{S}_{max}$. The weight of unevaporated spray may be computed as a function of $\bar{S}_{max}$ by considering a linear movement of the $\bar{S}$-axis to the right in Fig. 11, and assuming that all drops to the left of the origin have vanished. Fig. 12 shows the spray weight as a function of $\bar{S}_{max}$. Using a psychrometric chart and allowing for the adiabatic saturation, one can compute easily the change in wet-bulb depression with spray weight or $\bar{S}_{max}$ as shown in Fig. 13 for two initial conditions. Since $\bar{S}_{max}$ follows Equation [27] and $(T_{wb} - T'_{db})$ is known as a function of $\bar{S}_{max}$, Equation [27] may be integrated graphically to show $\bar{S}_{max}$ as a function of time, as illustrated in Fig. 14. Knowing how $\bar{S}_{max}$ changes with time one can then read values off Fig. 11 and prepare graphs such as are presented in Fig. 15 and 16 which portray $G(r)$ and $(4/3)\pi r^3 G(r)$ which represents the volumetric size distribution function.

Authors' Closure

The authors thank Professor Tribus for presenting his numerical results illustrating a case of particular practical interest.

Heat Transfer in a Pipe With Turbulent Flow and Arbitrary Wall-Temperature Distribution[1]

By C. A. SLEICHER, JR.,[2] and M. TRIBUS[3]

The first three eigenvalues and constants for the problem of heat flow to a constant property fluid in established turbulent flow in a round pipe are presented for all important values of Reynolds and Prandtl Moduli. These results permit one to compute heat transfer from nonisothermal pipe walls. Comparisons with experiment are good for all fluids from oils to liquid metals. The thermal entry length and rate of heat transfer for low Prandtl Moduli are shown to depend markedly on the wall-temperature profile.

Nomenclature

The following nomenclature is used in the paper:

A_n = constant defined by Equation [7]
B = constant, axial temperature gradient, deg F/unit dimensionless length
C_n = constant defined by Equation [4] or [8]
D = pipe diameter, ft
f = $f(r_+)$, dimensionless velocity, u/u_{avg}
g = $g(r_+)$, dimensionless total thermal diffusivity, $1 + \Pr \epsilon_H/\nu$
g_c = conversion factor, 32.2 lb-mass ft/lb-force sec²
h = heat-transfer coefficient, $qD/k(t_w - t_{mm})$, Btu/hr sq ft deg F
k = thermal conductivity, Btu/hr ft deg F
L = length, ft
Nu = local Nusselt number, hD/k
Nu_a = asymptotic or fully developed Nusselt number
Pe = Peclet number, Re Pr
Pr = Prandtl number, $C_p \mu/k$
q = heat flux, Btu/hr sq ft
r = radial distance, ft
r_0 = pipe radius, ft
r_+ = r/r_0
Re = Reynolds number, $D\, u_{avg} \rho/\mu$
R_n = eigenfunction defined by Equation [5]
t = total temperature, deg F
t_{mm} = mixed-mean temperature, deg F
t_0 = inlet temperature, deg F
t_w = wall temperature, deg F
u = mean velocity at a point, fps
u^+ = dimensionless velocity, $u/\sqrt{g_c \tau_0/\rho}$
u_{avg} = average bulk velocity in pipe, fps
x = axial distance, ft
x_+ = $2x/\mathrm{Pe}\, D$
y = $r_0 - r$, distance from wall, ft
y^+ = dimensionless distance from wall, $y\sqrt{g_c \tau_0/\rho}/\nu$
α = ϵ_H/ϵ_M, ratio of eddy diffusivities
ϵ_H = eddy diffusivity for heat, ft²/sec
ϵ_M = eddy diffusivity for momentum, ft²/sec
θ = dimensionless temperature, $(t - t_w)/(t_0 - t_w)$
λ_n = eigenvalue in Equation [5]
μ = viscosity, lb/ft sec
ν = kinematic viscosity, ft²/sec
ρ = density, pcf
τ_0 = shear stress at the wall, lb-force/sq ft

Introduction and Statement of Problem

The rate of heat transfer to fluids flowing in pipes is of great commercial interest and has been the object of hundreds of investigations over the past 80 years. These are admirably reviewed by McAdams (1).[4] In these investigations relatively little attention has been directed to the effect of wall-temperature distribution. Such effects are often small, but there are many cases of interest in which they are appreciable. The object of this paper is to provide an analysis which will yield the local rate of heat transfer to a fluid flowing turbulently in a pipe in which the wall temperature or heat flux varies in an arbitrary fashion.

Definition of System

This analysis is limited to smooth, straight pipes within which the fluid flow is fully developed. For this case the energy equation can be written

$$u \frac{\partial t}{\partial x} = \frac{1}{r} \frac{\partial}{\partial r} \left[r \left(\frac{\nu}{\Pr} + \epsilon_H \right) \frac{\partial t}{\partial r} \right] \quad \ldots\ldots [1]$$

The system satisfying this equation is subject to the following restrictions:

1 Fluid properties are constant.
2 Mean velocity in axial direction is independent of angular position.
3 Mean radial velocity is zero.
4 Mean temperature at any radius does not vary with time or angular position.
5 Frictional dissipation of energy is negligible.
6 Axial diffusion is negligible with respect to bulk transport in the x-direction.

The last restriction has been shown by Schneider (2) to introduce a negligible error for $\mathrm{Pe} > 100$.

Two wall-boundary conditions are of particular interest, uniform wall temperature and uniform wall-heat flux. The case of

[1] Investigation was conducted while the authors were at the University of Michigan, Department of Chemical and Metallurgical Engineering.

[2] Shell Development Company, Emeryville, Calif.

[3] Associate Professor, University of California, Los Angeles, Calif. Mem. ASME.

Contributed by the Heat Transfer Division of The American Society of Mechanical Engineers and presented at the Heat Transfer and Fluid Mechanics Institute, Stanford University, Calif., June 1956.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, May 11, 1956.

[4] Numbers in parentheses refer to the Bibliography at the end of the paper.

uniform wall temperature is considered here because the "step-function" solution can easily be used to solve the case of an arbitrary wall-temperature distribution by superposition (3). Reference (3) also shows how to use the uniform-wall-temperature solution to solve the case of arbitrary wall-heat flux. The boundary conditions for the "step function" are

$$t(x, r) = t_0,\ x < 0;\quad t(x, r_0) = t_w,\ x > 0;$$
$$t(0, r) = t_0,\quad r \neq r_0 \ldots\ldots [2]$$

Related Systems

Graetz, as reported by Jakob (4), first solved Equation [1] for laminar flow. Latzko (5) gave an approximate solution for turbulent flow with Pr = 1. Poppendiek and Harrison (6) review four pipe solutions limited to very low Peclet numbers. Other systems involving heat transfer from various surfaces with arbitrary temperature distribution are reviewed by Tribus and Klein (3, 7).

Form of the Solution

To solve Equation [1] with the Boundary Conditions [2], it is convenient to render it dimensionless by use of the following definitions

$$r_* = \frac{r}{r_0} \qquad f(r_*) = \frac{u}{u_{avg}} \qquad x_* = \frac{2x}{\mathrm{Re\ Pr}\ D}$$

$$\theta(x_*, r_*) = \frac{t - t_w}{t_0 - t_w} \qquad g(r_*) = \frac{\nu/\mathrm{Pr} + \epsilon_H}{\nu/\mathrm{Pr}}$$

Substitution of the foregoing values into Equation [1] yields

$$f\frac{\partial\theta}{\partial x_*} = \frac{2}{x_*}\frac{\partial}{\partial r_*}\left[r_* g \frac{\partial\theta}{\partial r_*}\right] \ldots\ldots [3]$$

with boundary conditions

$$\theta(x_*, r_*) = 1, x_* < 0;\ \theta(x_*, 1) = 0, x_* > 0;\ \theta(0, r_*) = 1, r_* \neq 1$$

By the usual method of separation of variables, the solution is

$$\theta = \sum_{n=0}^{\infty} C_n R_n \exp(-\lambda_n^2 x_*) \ldots\ldots [4]$$

in which $R_n(r_*)$ satisfies

$$\frac{d}{dr_*}[r_* g R_n'] + \lambda_n^2 \frac{f}{2} r_* R_n = 0 \ldots\ldots [5]$$

with the boundary conditions, $R_n(1) = 0$, $R_n(0) = 1$.

The heat flux to the fluid at the wall is given by

$$q(x) = +k\left(\frac{\partial t}{\partial r}\right)_{r=0} = \frac{-4k(t_0 - t_w)}{D} \Sigma A_n \exp(-\lambda_n^2 x_*) \ldots [6]$$

in which

$$A_n = -\frac{C_n R_n'(1)}{2} \ldots\ldots [7]$$

The equations are presented in the foregoing form in order to agree with the laminar flow case in Jakob (4) and Sellars, Tribus, and Klein (8).

Equation [5] with its boundary conditions belongs to the well-known class of differential equations called Sturm-Liouville systems (9). From the orthogonality property of the solutions, it can be shown (9) that the coefficients C_n are given by

$$C_n = \frac{\int_0^1 f r_* R_n dr_*}{\int_0^1 f r_* R_n^2 dr_*} \ldots\ldots [8]$$

The forms of $f(r_*)$ and $g(r_*)$ were obtained experimentally (10, 11). The λ_n and R_n were found with the aid of an electronic analog computer in which $f(r_*)$ and $g(r_*)$ were formed by a function generator.

The Velocity Function, $f(r_*)$

Measurements of turbulent-velocity distributions in pipes have been reported by many authors and are reviewed in reference (10). The results of these investigations are not in complete agreement, but the disagreements are not sufficient to cause enough difference in $f(r_*)$ to affect the solution of Equation [5] to a significant degree.

The Eddy Diffusivity Function, $g(r_*)$

The eddy diffusivity for momentum has been shown by von Karman (12), for example, to be related to the velocity in a pipe by

$$\epsilon_M = \frac{\left(\frac{\tau_0}{\rho}\right)\frac{r}{r_0}}{\frac{\partial u}{\partial y}} - \nu \ldots\ldots [9]$$

Also, it is customary to define

$$\alpha = \epsilon_H/\epsilon_M \ldots\ldots [10]$$

Thus one can write

$$g(r_*) = 1 + \alpha\ \mathrm{Pr}\left[\frac{\left(\frac{\partial u}{\partial y}\right)_{y=0}^{r_*}}{\left(\frac{\partial u}{\partial y}\right)} - 1\right] \ldots\ldots [11]$$

From Equation [11], $g(r_*)$ can be calculated if an accurate velocity distribution is available and if α is known.

Reynolds (13) first suggested that $\alpha = 1$, a statement of the idea that heat and momentum are transferred by the same mechanism. Jenkins (14) proposed a modification of Prandtl's mixing-length theory (15) by supposing that an eddy loses some of its heat and momentum as it travels from one layer of fluid to another of different temperature and velocity. Jenkins' analysis predicts that α will approach unity as turbulence increases, a trend that is clearly indicated by the experimental work of Isakoff and Drew (16), Corcoran, et al. (17), Brown, Amstead, and Short (18), and Sleicher (10). The analysis also predicts that α should decrease with decreasing Prandtl number. This trend is indicated by a comparison of the results of Isakoff and Drew or Brown, Amstead, and Short for mercury with those of Sleicher for air, all in pipe flow. The predicted low value of α for low Prandtl numbers agrees with the fact that most experimental values of the Nusselt number for liquid metals are below the predictions of Martinelli (19) and Lyon (20) based on $\alpha = 1$.

Although Jenkins' analysis predicts the right trends for α, the absolute values are lower than the experimental results of references (10, 16, 17, 18, 21). In view of the foregoing, the values of α used to determine $g(r_*)$ were calculated by multiplying Jenkins' value for any Prandtl number by a factor such that agreement was reached with the experimental results of Sleicher (10) for air. That is

$$\alpha\left(\frac{\epsilon_M}{\nu}, \mathrm{Pr}\right) = \frac{\alpha_S\left(\frac{\epsilon_M}{\nu}, \mathrm{air}\right)}{\alpha_J\left(\frac{\epsilon_M}{\nu}, \mathrm{air}\right)} \alpha_J\left(\frac{\epsilon_M}{\nu}, \mathrm{Pr}\right) \ldots\ldots [12]$$

where α_J is determined from Jenkins (analysis) and α_S from Sleicher (measurements).

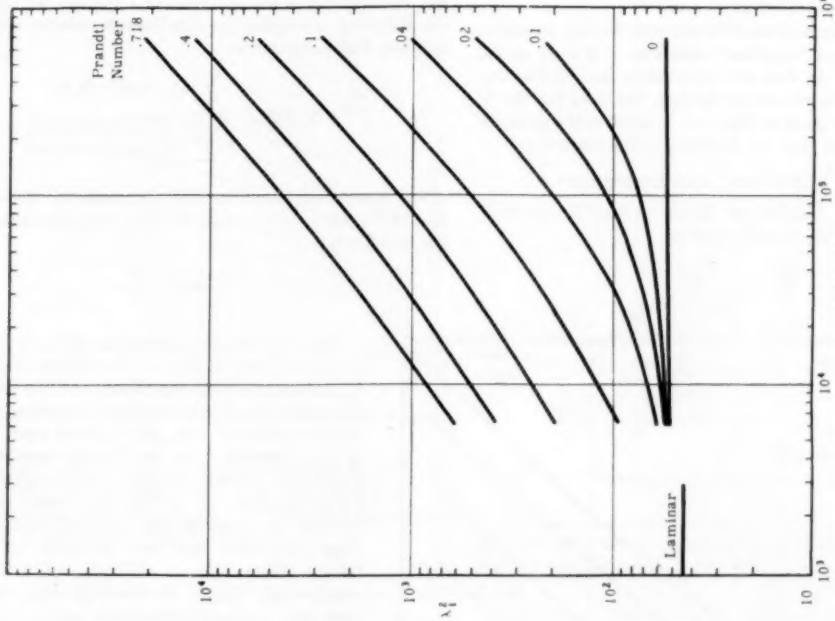


FIG. 2 SQUARE OF SECOND EIGENVALUE, λ_1^2, VERSUS REYNOLDS NUMBER

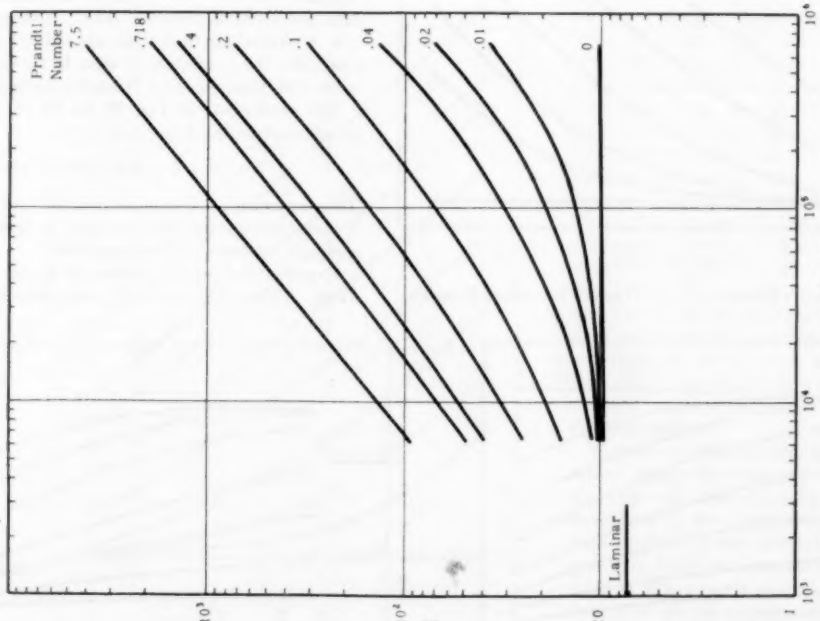


FIG. 1 SQUARE OF FIRST EIGENVALUE, λ_0^2, VERSUS REYNOLDS NUMBER

Analog-Computer Results

The analog computer was used to generate known solutions, such as the Graetz and "slug-flow" solutions and gave results within 2 per cent for the first few eigenvalues and coefficients. Smoothed and interpolated curves through the data for the 38 turbulent cases run are given in Figs. 1–9. Tables of the data and eigenfunctions for these runs can be found in reference (11).

Comparison of Results With Experiment

Asymptotic Nusselt Number for Uniform Wall-Temperature. From the definition of the Nusselt number

$$\mathrm{Nu} = \frac{hD}{k} = \frac{qD}{k(t_w - t_m)}$$

and by appropriate manipulation of Equation [4], one can derive the following expression for the Nusselt number for the case of uniform wall temperature

$$\mathrm{Nu}(x) = \frac{\sum_n A_n \exp(-\lambda_n^2 x_*)}{2\sum_n \frac{A_n}{\lambda_n^2}\exp(-\lambda_n^2 x_*)} \quad \ldots\ldots [13]$$

Far downstream from the thermal entrance, all terms but the first of Equation [13] become small so that the asymptotic Nusselt number is

$$\mathrm{Nu}_a = \frac{\lambda_0^2}{2} \quad \ldots\ldots [14]$$

Thus the Nusselt number for this case is simply half of the λ_0^2 given in Fig. 1. For air the Nusselt number is in agreement with experiment and the Dittus-Boelter Equation [1], which was to be expected since experimental values of u/u_m and ϵ_H were used. The region of most interest is at low Prandtl numbers. In this region the Nusselt number is often correlated against Peclet number, as in Fig. 10, which shows a line through the heating data of Gilliland, Musser, and Page (22) which has been corrected for thermal-entrance effects. These authors used a uniform wall-temperature system for their heating runs, but the tube had a length-to-diameter ratio of 45, which is sufficiently short that the average Nusselt number was 5 to 35 per cent higher than the asymptotic value. The modification to their data was made by multiplying their data by $\mathrm{Nu}_a/\mathrm{Nu}_{avg}$ as determined from the data and methods presented here. Although the methods of evaluating their data are subject to inaccuracies (23), the prediction is seen to be in good agreement with their modified Nusselt numbers.

The predictions of Fig. 10 for $\mathrm{Pr} < 0.05$ may be represented within 2 per cent by

$$\mathrm{Nu}_a = 4.8 + 0.0150\ \mathrm{Pe}^{0.91}\ \mathrm{Pr}^{0.30} \quad \ldots\ldots [15]$$

This equation is recommended for the asymptotic Nusselt number for heat transfer to liquid metals in a pipe at uniform wall temperature.

Asymptotic Nusselt Number for Uniform Wall-Heat Flux. Tribus and Klein (3) have shown how to use

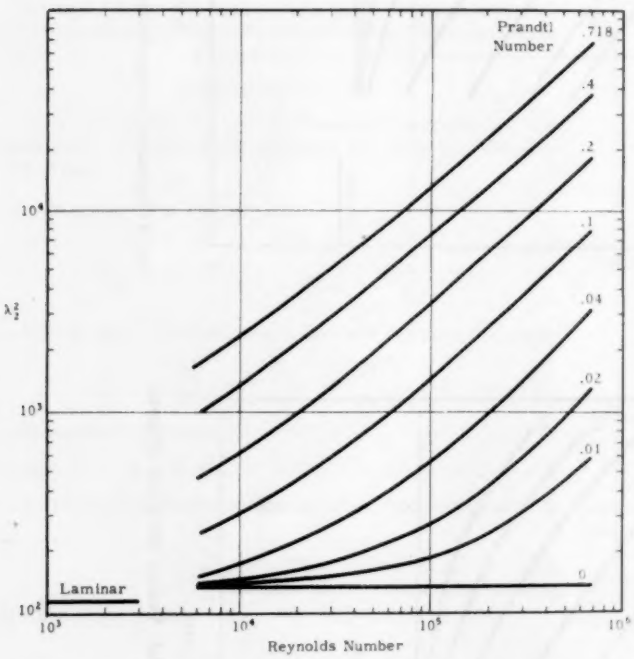


Fig. 3 Square of Third Eigenvalue, λ_2^2, Versus Reynolds Number

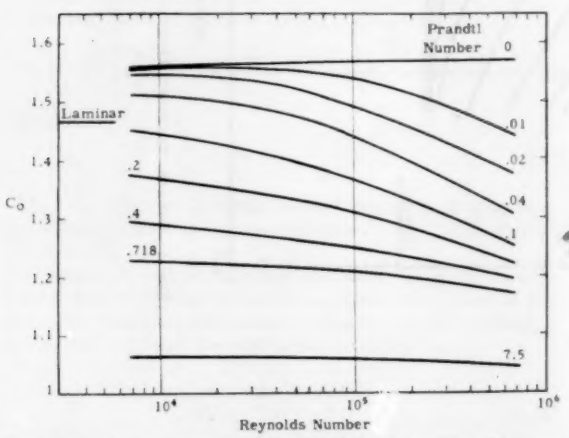


Fig. 4 C_0, Constant in Equation [4] Versus Reynolds Number

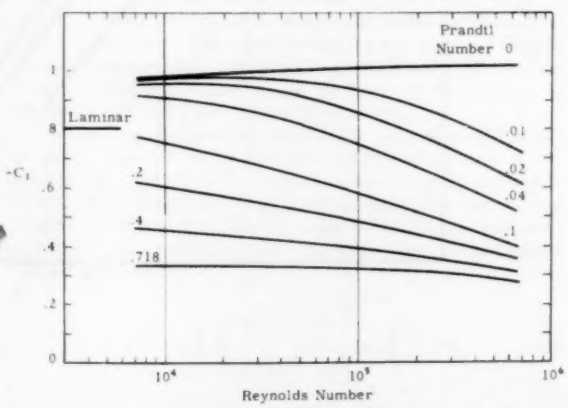


Fig. 5 $-C_1$, Constant in Equation [4] Versus Reynolds Number

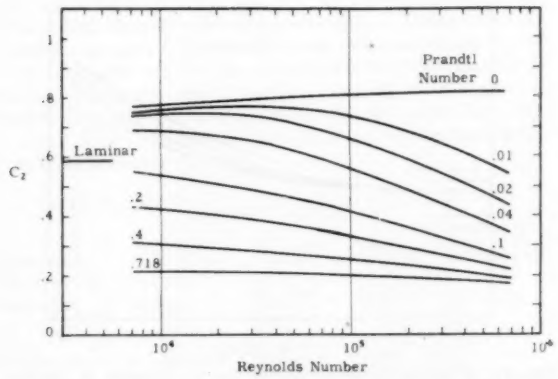


FIG. 6 C_2, CONSTANT IN EQUATION [4] VERSUS REYNOLDS NUMBER

the solution for uniform wall-temperature systems to solve the case of arbitrary wall temperature or heat flux. With the use of their techniques, it can be shown (11) that the asymptotic Nusselt number for uniform wall-heat flux is

$$\mathrm{Nu}_a = \frac{1}{16 \sum \frac{A_n}{\lambda_n^4}} \quad \ldots\ldots\ldots\ldots\ldots [16]$$

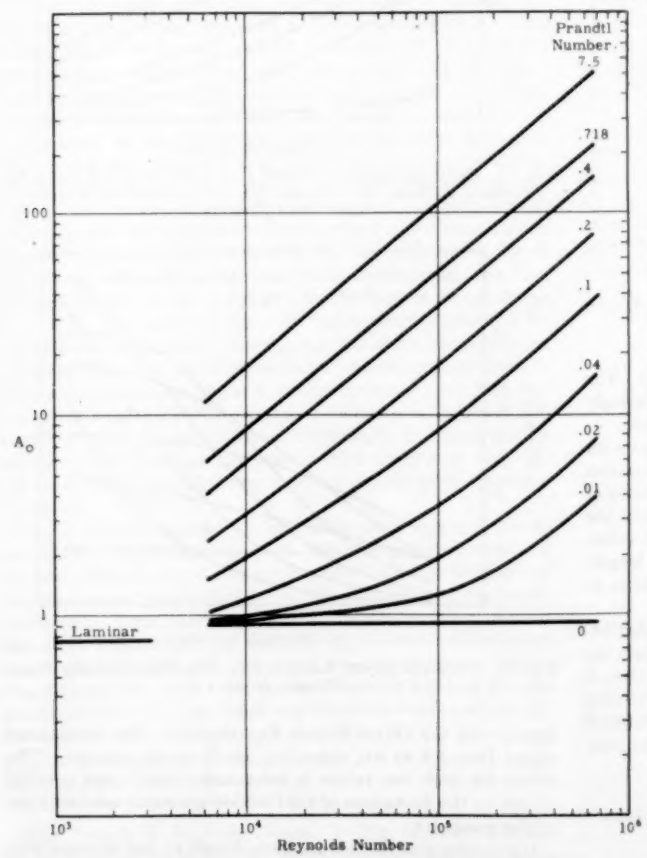


FIG. 7 A_0, CONSTANT IN EQUATION [6] VERSUS REYNOLDS NUMBER

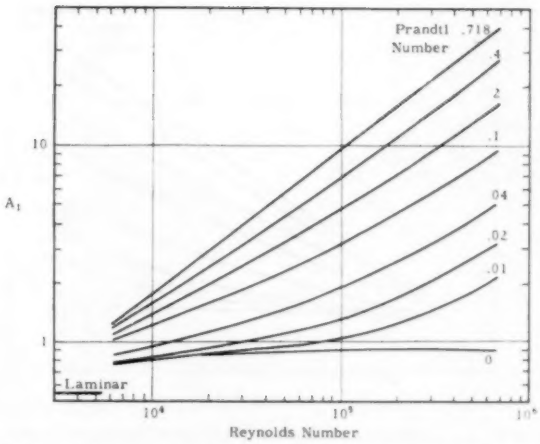


FIG. 8 A_1, CONSTANT IN EQUATION [6] VERSUS REYNOLDS NUMBER

This series converges extremely rapidly. For example, for the laminar case, which converges more slowly than the turbulent, the first two terms give 4.379 whereas the exact value is 48/11 = 4.364.

The dashed lines of Fig. 11 are plots of Equation [16] versus Peclet number for the liquid-metal region. Shown also are mean lines through the data of several experimental investigations. Lubarsky and Kaufman (23), in a report dated March, 1955, analyzed very thoroughly all reported experimental work on liquid-metal heat transfer. They report sources of error in many of the experimental procedures or data analysis, and reports containing such faults were not considered for comparison here. In addition, some reports were not considered for one or more of the following reasons: L/D ratio so short that the asymptotic value of Nu was not achieved, velocity profile not developed, rough-walled pipe, presence of probe disturbed flow. This culling left seven references, whose location can be found in Lubarsky and Kaufman. The lines shown for English and Barret; Johnson, Hartnett, and Clabaugh; Johnson, Clabaugh, and Hartnett; and Stromquist are from their unchanged data. The lines for Styrikovich and Semenovker, and Trefethen are shown as re-evaluated by Lubarsky and Kaufman (23). The Seban data are shown as points for the three runs taken immediately after cleaning of the tube walls.

The agreement with the experimental results is excellent, especially when one considers that at some of the very low Peclet numbers shown the flow may have been in the laminar or transition region. All these results fall well below the theoretical line proposed by Lyon and also shown in Fig. 11. The results of the present analysis shown in Fig. 11 for $\mathrm{Pr} < 0.05$ can be represented within 2 per cent by

$$\mathrm{Nu} = 6.3 + 0.0160\ \mathrm{Pe}^{0.91}\ \mathrm{Pr}^{0.30} \quad \ldots\ldots [17]$$

This equation is recommended for the asymptotic Nusselt number for liquid metals in a pipe at uniform heat flux.

The ratio of the asymptotic Nusselt number at uniform heat flux to that at uniform wall temperature is shown in Fig. 12 and confirms the similar but

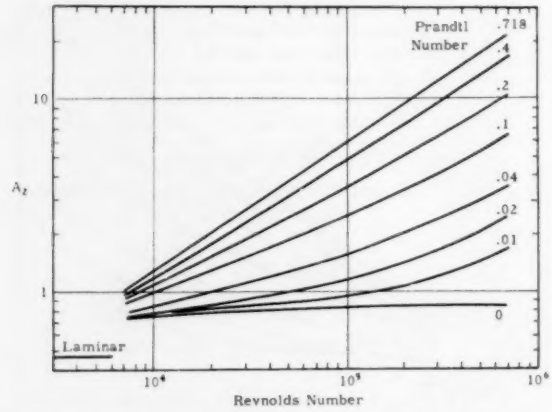


FIG. 9 A_3, CONSTANT IN EQUATION [6] VERSUS REYNOLDS NUMBER

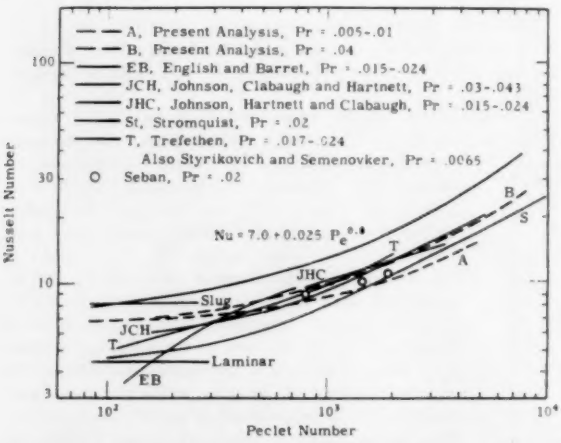


FIG. 11 NUSSELT NUMBER VERSUS PECLET NUMBER IN LIQUID METAL REGION FOR UNIFORM HEAT FLUX

more limited calculations of Seban and Shimazaki (24). The figure illustrates that the wall-temperature distribution strongly affects the heat-transfer coefficient at low Peclet numbers.

Thermal-Entry Length. Fig. 13 shows thermal-entry lengths for a pipe at constant wall temperature, calculated from Equation [13]. The thermal-entry length is defined here as the number of diameters downstream from the beginning of heating at which the Nusselt number reaches 2 per cent of its asymptotic value. Other authors have used a 1 per cent criterion for entry length, but reported experimental results are not sufficiently precise to warrant use of the 1 per cent definition for comparison.

The line for water (Prandtl number 7.5) is shown dotted because there is some doubt of its accuracy. In order to obtain the correct value for the asymptotic Nusselt number for water, it was necessary to use values of the eddy diffusivity much lower than expected. In calculating the eddy diffusivity for water, it was assumed that $\epsilon_H = \epsilon_M$ in the central portion of the pipe. Near the wall the equation of Deissler (25)

$$\frac{\epsilon_H}{\nu} = \alpha \frac{\epsilon_M}{\nu} = 0.0154\alpha u^+ y^+ \left[1 - \exp\left(-0.0154\, u^+ y^+\right)\right]$$

was used, and α was varied until the asymptotic Nusselt number

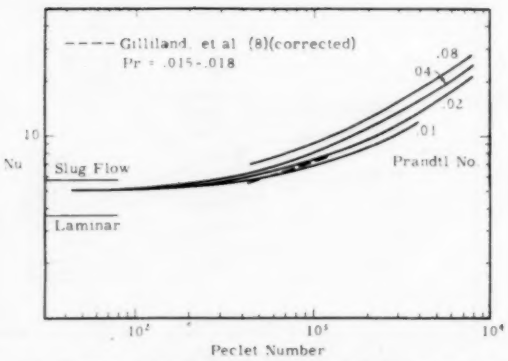


FIG. 10 NUSSELT NUMBER VERSUS PECLET NUMBER IN LIQUID-METAL REGION FOR UNIFORM WALL TEMPERATURE

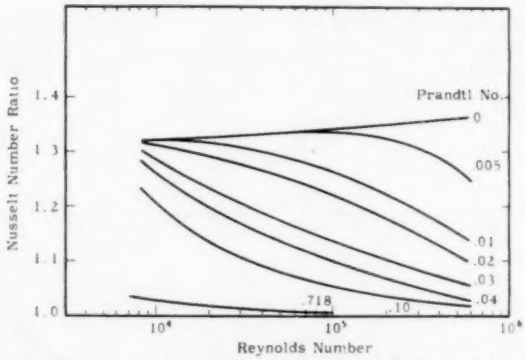


FIG. 12 RATIO OF NUSSELT NUMBER AT UNIFORM HEAT FLUX TO NUSSELT NUMBER AT UNIFORM WALL TEMPERATURE VERSUS REYNOLDS NUMBER

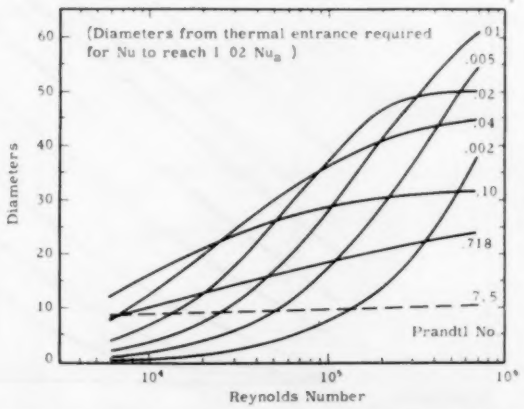


FIG. 13 THERMAL ENTRY LENGTH FOR A PIPE AT UNIFORM WALL TEMPERATURE

agreed with the Dittus-Boelter Equation [1]. The values used varied from 0.4 to 0.6, depending on Reynolds number. The reason for such low values is not certain, but it was possibly caused by the limitations of the function generator used with the analog computer.

It is curious that the thermal-entry length at first increases with increasing Prandtl number and then begins to decrease. This

variation is contrary to the prediction of Berry (26), but his results at high Prandtl number are considerably higher than the experimental entry lengths of Hartnett (27) for oil and water. Both Hartnett's data and the predictions of Levy (28) and Berry indicate that, as the Prandtl number increases above about 10, there is very little increase in entry length at a given Reynolds number. Thus the dotted line in Fig. 13 is the prediction for all Prandtl numbers above 7.5.

TABLE 1 SUMMARY OF THERMAL-ENTRY-LENGTH INVESTIGATIONS

Investigator	Type of investigation	Boundary condition	Prandtl number	Thermal-entry length, x/D Re = 10^4	Re = 10^5
Present results..	Analytical	Uniform wall temp	7.5	9	11
Deissler (31)...	Analytical	Uniform heat flux	10	2–3	2–3
Berry (26)[a]....	Analytical	Uniform wall temp	10	13	17
Hartnett (27)...	Experimental	Uniform heat flux	7–200	10	15
Present results..	Analytical	Uniform wall temp	0.718	10	19
Deissler (31)...	Analytical	Uniform wall temp	0.73	2	7
Deissler (31)...	Analytical	Uniform heat flux	0.73	3	7
Latzko (5).....	Analytical	Uniform wall temp	1	6	11
Berry[a] (26)....	Analytical	Uniform wall temp	1	12	17
Boelter et al. (29) (extrapolated)	Experimental	Uniform wall temp	0.72	8	15

[a] Berry used a 1 per cent instead of a 2 per cent entry-length criterion. A factor of about 0.7 would convert his results to a 2 per cent entry length.

Table 1 summarizes the results of previous investigations and compares them to the present at two Reynolds numbers. This table is similar to the one reported by Hartnett except that a 2 per cent entry length is used here, whereas Hartnett apparently used 1 per cent in determining the entry lengths of Boelter, Latzko, and Deissler.

The entry lengths calculated in this investigation are somewhat higher than the analytical results of Latzko and Deissler. The experimental results of Boelter (29) fall about halfway between Latzko's analysis and the present and are considerably above the results of Deissler. For higher Prandtl numbers the present results agree very well with the data of Hartnett for oil and water, which were, however, taken at uniform heat flux. This agreement and the fact that the entry lengths for air at uniform heat flux and uniform wall temperature calculated by Deissler are nearly equal indicate that the entry lengths for the two cases are in fact about the same for Prandtl numbers above 1. This observation disagrees with the prediction of Levy (28) that the entry lengths for the two cases differ by a factor of 3.6. Levy's analysis is oversimplified, however, in that it assumes a sublayer of a calculated thickness in which there is no eddy diffusion bounded by a well-stirred fluid in which the eddy diffusivity is infinite.

At low Prandtl numbers the calculated entry length is in general agreement with Johnson, et al. (30), who estimate from their experimental data that for Prandtl numbers of 0.020 to 0.045 and Reynolds numbers of 10,000 to 100,000 the thermal-entry length is about 30. Their data were taken at uniform heat flux. In this region the uniform heat flux and uniform wall-temperature entry lengths probably differ more but not much more than at higher Prandtl numbers. Entry-length calculations for a fluid of Prandtl number 0.025 flowing in a pipe at a Reynolds number of 120,000 give the following results:

Entry length at uniform wall-temperature	40
Entry length at uniform heat flux	44
Entry length at linear wall-temperature	195

Heat Transfer in the Entry Region. In Fig. 14 are plotted the experimental values of Nu/Nu_a at two Reynolds numbers taken from Boelter, et al. (29), and the corresponding line calculated by Equation [13]. There is some scatter in the data, but the line follows the data fairly well out to about $x/D = 4$. For the region closer than that, more eigenvalues and constants are needed for the case of air. That close to the entrance the boundary layer calculations of Deissler (31) undoubtedly give more accurate results.

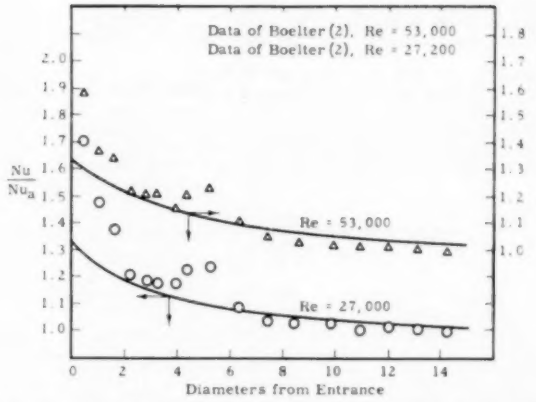


FIG. 14 HEAT TRANSFER TO AIR IN THE THERMAL ENTRANCE REGION

Several authors report experimental Nusselt numbers in the thermal-entrance region for uniform heat flux. These are not compared here because more than three eigenvalues are desirable for calculating heat transfer for those cases. Methods for handling such conditions are under consideration.

SUMMARY OF EQUATIONS FOR ESTIMATING HEAT TRANSFER

Below are listed equations for computing the temperature distribution, rate of heat transfer, mixed-mean temperature, and Nusselt number for three wall-temperature conditions. With the aid of these equations and the constants given in Figs. 1–9, it is a simple matter to calculate heat transfer in a pipe for any of the three cases.

Uniform Wall Temperature

$$\frac{t - t_w}{t_0 - t_w} = \sum_n C_n R_n \exp(-\lambda_n^2 x_*) \quad [18]$$

$$q(x_*) = \frac{-4k(t_0 - t_w)}{D} \sum_n A_n \exp(-\lambda_n^2 x_*) \quad [19]$$

$$t_m = t_w - 8(t_w - t_0) \sum_n \frac{A_n}{\lambda_n^2} \exp(-\lambda_n^2 x_*) \quad [20]$$

$$\mathrm{Nu}(x_*) = \frac{\sum_n A_n \exp(-\lambda_n^2 x_*)}{2 \sum_n \frac{A_n}{\lambda_n^2} \exp(-\lambda_n^2 x_*)} \quad [21]$$

After only the first exponential is important, Equation [21] reduces to the asymptotic Nusselt number

$$\mathrm{Nu}_a = \frac{\lambda_0^2}{2} \quad [22]$$

Linear Wall Temperature. If

$$t_w(x_*) - t_0 = Bx_* \quad [23]$$

$$t(x_*, r_*) - t_0 = Bx_* - B\sum_n \frac{C_n R_n}{\lambda_n^2} + B\sum_n \frac{C_n R_n}{\lambda_n^2} \exp(-\gamma_n^2 x_*) \quad [24]$$

$$q(x_*) = \frac{Bk}{2D} + \frac{4Bk}{D}\sum_n \frac{A_n}{\lambda_n^2} \exp(-\lambda_n^2 x_*) \quad [25]$$

$$t_m - t_0 = Bx_* + 8B\sum_n \frac{A_n}{\lambda_n^4}[\exp(-\lambda_n^2 x_*) - 1] \quad [26]$$

$$\mathrm{Nu}(x) = \frac{1 + 8\sum_n \frac{A_n}{\lambda_n^2} \exp(-\lambda_n^2 x_*)}{16\sum_n \frac{A_n}{\lambda_n^4}[1 - \exp(-\lambda_n^2 x_*)]} \quad [27]$$

For the asymptotic Nusselt number, Equation [27] reduces to

$$\mathrm{Nu}_\infty = \frac{1}{16\sum_n \frac{A_n}{\lambda_n^4}} \quad [28]$$

Uniform Heat Flux at the Wall. The method of deriving these equations is explained by Tribus and Klein (3) and Sellars, Tribus, and Klein (8). Those references should be consulted for further details. Let

$$H(s) = 2\sum_n \frac{A_n}{s + \lambda_n^2} \quad [29]$$

Then

$$H'(s) = -2\sum_n \frac{A_n}{(s + \lambda_n^2)^2} \quad [30]$$

Now let γ_m^2 be the values satisfying $H(-\gamma_m^2) = 0$. The temperature distribution is then given by

$$t(x_*, r_*) - t_0 = \frac{qD}{2k}\left[4x_* + 32\sum_n \frac{A_n}{\lambda_n^4} + \sum_m \frac{\exp(-\gamma_m^2 x_*)}{\gamma_m^4 H'(-\gamma_m^2)} - \sum_n C_n R_n \sum_m \frac{1 - \exp(-\gamma_m^2 x_*)}{\gamma_m^2(\lambda_n^2 - \gamma_m^2)}\right] \quad [31]$$

The first summation in the brackets is equal to

$$-\sum_m \frac{1}{\gamma_m^4 H'(-\gamma_m^2)}$$

given by Sellars, Tribus, and Klein. It converges much faster than their expression in the foregoing, however.

Remaining expressions of importance for uniform heat flux are

$$t_m - t_0 = \frac{2qD}{k} x_* \quad [32]$$

$$\mathrm{Nu}(x_*) = \frac{1}{16\sum_n \frac{A_n}{\lambda_n^4} + \frac{1}{2}\sum_m \frac{\exp(-\gamma_m^2 x_*)}{\gamma_m^4 H'(-\gamma_m^2)}} \quad [33]$$

Acknowledgment

The authors express their appreciation to Profs. Donald L. Katz, Stuart W. Churchill, and John R. Sellars of the University of Michigan for their suggestions and criticisms throughout the investigation. Gratitude is also extended to the Department of Aeronautical Engineering for the use of its analog-computer equipment.

BIBLIOGRAPHY

1 "Heat Transmission," by W. H. McAdams, third edition, McGraw-Hill Book Co., Inc., New York, N. Y., 1954, pp. 205–251.

2 "Effect of Axial Fluid Conduction on Heat Transfer in the Entrance Regions of Parallel Plates and Tubes," by P. J. Schneider, published in this issue, pp. 765–773.

3 "Forced Convection From Nonisothermal Surfaces," by M. Tribus and J. Klein, in "Heat Transfer, a Symposium held at the University of Michigan during the Summer of 1952," Engineering Research Institute, University of Michigan, Ann Arbor, Mich., 1953, pp. 211–235.

4 "Heat Transfer," by M. Jakob, John Wiley and Sons, Inc., New York, N. Y., vol. 1, 1949, pp. 451–464.

5 "Heat Transfer in a Turbulent Liquid or Gas Stream," by H. Latzko, NACA TM 1068, 1944; original in *Zeitschrift für angewandte Mathematik und Mechanik*, vol. 1, no. 4, 1921.

6 "Remarks on Thermal Entrance-Region Heat Transfer in Liquid-Metal Systems," by H. F. Poppendiek and W. B. Harrison, paper presented at Heat Transfer Symposium in St. Louis, Mo., Dec. 13–16, 1953.

7 "Forced Convection Through a Laminar Boundary Layer Over an Arbitrary Surface With an Arbitrary Temperature Variation," by M. Tribus and J. Klein, *Journal of the Aeronautical Sciences*, vol. 22, 1955, pp. 62–64.

8 "Heat Transfer to Laminar Flow in a Round Tube or Flat Conduit, The Graetz Problem Extended," by J. R. Sellars, M. Tribus, and J. Klein, Trans. ASME, vol. 78, 1956, pp. 441–448.

9 "Modern Operational Mathematics in Engineering," by R. V. Churchill, McGraw-Hill Book Co., Inc., New York, N. Y., 1944, pp. 242–266.

10 "Experimental Velocity and Temperature Profiles for Air in Turbulent Flow in a Pipe," by C. A. Sleicher, Jr., paper submitted to ASME, 1956.

11 "Heat Transfer in a Pipe With Turbulent Flow and Arbitrary Wall-Temperature Distribution," by C. A. Sleicher, Jr., PhD thesis in Chemical Engineering, University of Michigan, Ann Arbor, Mich., 1955. Available on microfilm from University Microfilms, Ann Arbor, Mich.

12 "The Analogy Between Fluid Friction and Heat Transfer," by Th. von Karman, Trans. ASME, vol. 61, 1939, pp. 705–710.

13 "On the Extent and Action of the Heating Surface for Steam Boilers," by O. Reynolds, Proceedings of the Manchester Literary and Philosophical Society, vol. 14, 1874, pp. 7.

14 "Variation of Eddy Conductivity With Prandtl Modulus and Its Use in Prediction of Turbulent Heat Transfer Coefficients," by R. Jenkins, Heat Transfer and Fluid Mechanics Institute Preprints, Stanford University Press, Stanford, Calif., 1951, pp. 147–158.

15 "Über die ausgebildete Turbulenz," by L. Prandtl, Proceedings of the Second International Congress for Applied Mechanics, Zurich, Switzerland, 1926, pp. 62–74.

16 "Heat and Momentum Transfer in Turbulent Flow of Mercury," by S. E. Isakoff and T. B. Drew, General Discussion on Heat Transfer, The Institution of Mechanical Engineers, London, England, 1951, pp. 405–409. Also, PhD thesis of Isakoff, University Microfilms, Ann Arobr, Mich., Microfilm No. 4200.

17 "Temperature Measurements in Turbulent Gas Streams," by W. H. Corcoran, F. Page, Jr., W. G. Schlinger, and B. H. Sage, *Industrial and Engineering Chemistry*, vol. 44, 1952, pp. 410–430.

18 "The Transfer of Heat and Momentum in a Turbulent Stream of Mercury," by H. E. Brown, B. H. Amstead, and B. E. Short, ASME Paper No. 55—A-106.

19 "Heat Transfer to Molten Metals," by R. C. Martinelli, Trans. ASME, vol. 69, 1947, pp. 947–959.

20 "Liquid Metal Heat-Transfer Coefficients," by R. N. Lyon, *Chemical Engineering Progress*, vol. 47, 1951, pp. 75–79.

21 "Heat Transfer Through Turbulent Friction Layers," by H. Reichardt, NACA TM 1047, 1943.

22 "Heat Transfer to Mercury," by E. R. Gilliland, R. J. Musser, and W. R. Page, General Discussion on Heat Transfer, The Institution of Mechanical Engineers, London, England, 1951, pp. 402–404.

23 "Review of Experimental Investigations of Liquid-Metal Heat Transfer," by B. Lubarsky and S. J. Kaufman, NACA TN 3336, 1955.

24 "Heat Transfer to a Fluid Flowing Turbulently in a Smooth Pipe With Walls at Constant Temperature," by R. A. Seban and T. T. Shimazaki, Trans. ASME, vol. 73, 1951, pp. 803–808.

25 "Analysis of Turbulent Heat Transfer, Mass Transfer, and Friction in Smooth Tubes at High Prandtl and Schmidt Numbers," by R. G. Deissler, NACA TN 3145, 1954.

26 "Non-Uniform Heat Transfer to Fluids Flowing in Conduits," by V. J. Berry, *Applied Scientific Research*, Section A, vol. 4, 1953, pp. 61–75.

27 "Experimental Determination of the Thermal Entrance Length for the Flow of Water and of Oil in Circular Pipes," by J. P. Hartnett, Trans. ASME, vol. 77, 1955, pp. 1211–1220.

28 "Heat-Conduction Methods in Forced-Convection Flow," by S. Levy, Trans. ASME, vol. 78, 1956, p. 1627.

29 "An Investigation of Aircraft Heaters—XXVII. Distribution of Heat Transfer Rate in Entrance Section of a Circular Tube," by L. K. M. Boelter, D. Young, and H. W. Iversen, NACA TN 1451, 1948.

30 "Heat Transfer of Molten Lead-Bismuth Eutectic in Turbulent Pipe Flow," by H. A. Johnson, J. P. Hartnett, and W. J. Clabaugh, 1952 Heat Transfer and Fluid Mechanics Institute, Preprints of Papers, Stanford University Press, Stanford, Calif., 1952, pp. 5–18.

31 "Turbulent Heat Transfer and Friction in the Entrance Regions of Smooth Passages," by R. G. Deissler, Trans. ASME, vol. 77, 1955, pp. 1221–1233.

Practical Equipment-Replacement Economics

By P. A. SCHEUBLE, JR.,[1] MILWAUKEE, WIS.

There is nothing new about the basic problems encountered when replacing manufacturing equipment since we have been doing it for thousands of years at an accelerating rate. Often, it does not take any fancy figuring to tell us that a replacement is required. The basic reasons for replacement have not changed, except that the pressure for change has increased under present-day competition and advancements in technology. We find ourselves confronted with larger capital expenditures to keep up with the drive on reducing unit cost and increasing quality. Furthermore, there is competition between various equipment proposals and other projects in their demands on capital funds. The author discusses various principles on which the economic replacement of equipment must be based.

Introduction

IT HAS long been recognized that decisions on equipment purchases can have a profound effect on the future course of the business and its financial stability. Many managements, while closely controlling the budgeting and expenditure of capital funds at the top-executive level, have no definite policy or procedure for determining if and when a replacement should take place. It is thought that a sound approach to this problem can do much to stimulate thinking throughout the organization, about how to get the most for our investment dollar. What is needed is a method of analysis that will provide simple and understandable ground rules, so that various proposals can be compared in regard to capital availability, and that meritorious proposals can stand on their own feet, or at least have an adequate explanation of why they are not accepted.

We must avoid oversophistication of theory and keep the language simple, since the initiative for many changes comes from operating units in the organization, as well as the engineering departments. However, the method must relate realistically, the physical and economic factors involved. At the same time the method must be flexible enough to permit the evaluation of estimates and assumptions in relation to the possible range of error. The approach must be an aid to judgment, not a formula that turns out a "yes" or "no" answer.

Equipment-replacement problems have been analyzed from various viewpoints, ranging from the "one-horse shay" attitude of not replacing until the present equipment falls apart, to that of replacing simply because a newer piece of equipment may indicate an apparent operating cost saving over the one in use. They have varied from mandatory payback requirements as short as 1 to 2 years, to concepts primarily concerned with accounting and tax lives. But these diverse views are not answers in themselves; they are only symptomatic of the problems of equipment replacement.

[1] Assistant to Vice-President-Manufacturing, A. O. Smith Corporation, Milwaukee, Wis. Assoc. Mem. ASME.

Contributed by the Management Division and presented at the Engineering Management Conference, St. Louis, Mo., March 14–15, 1956, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, November 16, 1954. Paper No. 56—MGT-2.

Items to Consider

A replacement decision usually involves much more than a consideration of comparative annual operating costs and the effect of equipment wear. There is still the problem of determining the most advantageous use of available capital funds and making sure that a desired return on the invested capital will be received. This problem, in turn, is inextricably bound up with making estimates of the future as to:

1 Whether the necessary product volume will be maintained to realize the operating-cost advantages of the new equipment.

2 Whether overhaul of present equipment would be sufficient, and for how long.

3 Whether there are possible alternative uses of that equipment in the future.

4 Whether more advantageous equipment will appear on the market in the near future.

Also involved is the consideration of the possible adverse effects of not replacing; i.e., higher operating costs, lower quality, poorer customer service, forsaking of additional volume, loss of markets, and slowing up of long-range plans.

Although these factors may appear formidable at first glance, their evaluation, backed up by reasonable facts and figures, can be very effective in selling worth-while equipment ideas up the line. Of course, included in the consideration of equipment replacement are many factors often thought of as intangibles. The fact is, however, that such items, for instance, quality, can be measured quantitatively in terms of scrap cost, rework, inspection, and possible loss of sales volume. Other long-range considerations, while difficult to measure, can often be determined with workable accuracy by the use of cost and profit estimates (so long as their reasonableness is established and their lack of preciseness is taken into account in the interpretation of final results).

What Is Meant by Equipment Replacement?

What do we mean by equipment replacement? The first thing that comes to mind of course is replacing present equipment because of excessive wear or breakdown, or replacing because a much better type of equipment is available, or is to be designed. Actually, the term "equipment replacement" is rather broad in its application, since it also can mean, for instance, the replacement of one process by another. In fact, we limit our approach when we think only in terms of equipment and groups of equipment. The use of the term can cover any situation where unit costs can be reduced or profits increased by a capital expenditure. Obviously, when we replace we want to take advantage of the advances made in design and new concepts. On the other hand, sometimes overhaul of present equipment is adequate.

The degree of need for replacement varies from company to company according to sales volume, product character, competition, and so on. Thus, in a single product, high-volume operation, on the one hand, a small decrease in cost per unit may easily justify replacement. The management of a more diversified plant, on the other hand, may need to ponder longer before deciding on the replacement; here the question may be whether the new equipment will give the shop more versatility and whether a greater volume of business could be handled because of its advantages. Also entering into the problem of replacement

is the type of service the particular equipment is subject to and the type of maintenance it receives.

Need for Replacing Equipment

The need for replacing equipment may stem from any one or a combination of the following reasons:

1 Present equipment is not functioning properly, with one or more of these results: Excessive idle time, poor quality, increased labor costs, high operating costs, abnormal maintenance expense. Of course, overhaul may in many cases correct these conditions.

2 There is a current or an anticipated need for expanded or diversified capacity.

3 New or improved products have been introduced for competitive or other reasons.

4 More efficient equipment is available on the market, even though the present equipment may still perform satisfactorily.

5 Increased mechanization will reduce labor costs and/or increase quality. Reduction of fatigue is also a consideration.

6 Process and methods changes have been developed to reduce costs or to increase quality.

7 It is necessary to eliminate safety hazards; here the protection of workers is paramount, but there also may be gains in the form of increased output and reduction of industrial-compensation costs.

8 It is possible to achieve more efficient production and reduce costs by plant rearrangement, better material handling, establishment of manufacturing services, and so on.

There are various methods of equipment-replacement analysis. Those which incorporate the time value of money (or interest rates), it is believed, have the sounder approach since this concept cannot be avoided when considering the acquisition and use of money. The underlying relationship between money, the time cost of money, and its period of use, has been understood and used in financial circles and some types of engineering studies for many years. It should be emphasized that it is how these variables are used in relating physical and economic factors, which determines the value of a particular replacement-analysis method as a practical tool in which we can have confidence. It must tie in with experience and common sense.

Analyzing Proposals

More specifically, the method of analysis that is proposed measures the annual cost savings or profit potential of the proposed replacement against the capital investment made, taking into consideration its utilization, the desired percentage return on investment, the capital-recovery period, wear, and the service life of the equipment. This implies three basic tests which a replacement should pass before it is approved:

1 Its capital-recovery period must reasonably correspond with its period of potential utilization at a reasonable assumed average annual volume.

2 Its annual savings, or profit potential, after paying for its amortization and return on investment over its service life, must compete with other replacement proposals in their demand on capital funds.

3 Some consideration should be given to fitting in the proposal with the future plans of the company, especially in regard to equipment versatility and interchangeability.

As a practical matter, good judgment depends on standards of measurement. The essence of this approach is the use of eight yardsticks, which taken together, form the basis for a sound conclusion. They are not listed in order of importance but as they will be discussed.

1 The operating cost savings or increased-profit potential (or protection thereof) afforded by making the replacement.

2 The rate of return desired on the invested money.

3 The capital-recovery period, and its relation to the period of equipment utilization.

4 The service life of the equipment, and its relation to wear and obsolescence.

5 The total money invested in the replacement.

6 The net savings over the service life of the equipment.

7 The priority of a replacement in relation to other replacements and projects and the availability of capital funds.

8 The effect of taxes.

Annual Cost Savings. The usual starting point of a replacement analysis is the determination of the annual cost savings or dollar-profit improvement afforded by the new equipment, in comparison with the status quo. To do this we must have a reasonable idea as to potential volume. If present equipment can be overhauled, the cost-savings aspects of that alternative also must be determined. The study starts from now. What we have spent in the past is meaningful only if it makes a contribution to the future. It is a sunk cost. We thus have two basic replacement situations: (1) The present equipment can no longer do its job because of deterioration, now, or in the foreseeable future, and we must make the decision whether to spend further money on a major overhaul or replace it with the same or better equipment. A decision is relatively mandatory. (2) Factors other than wear or deterioration enter into the picture, such as obsolescence, and the cost-savings possibilities make replacement appear attractive.

Actually, both basic situations are present in varying degrees in almost any replacement problem.

Comparison of Operating Costs. The comparison of annual operating costs can embrace many factors, depending on the particular application. To mention a few: Labor costs, fringe labor benefits, maintenance and repair, scrap costs and rework costs, quality costs, setup time, down time, space, taxes and insurance, subcontract costs, and so forth. The time and effort needed to get figures for all of these factors may seem excessive. Yet the actual time spent is small in comparison with the capital outlay usually under consideration. The more limited a company's funds, the less it can afford a skin-deep review. Of course, judgment must play a part in developing such cost figures, and some objections may be raised against the crystal-ball aspect of trying to predict future volume, but there is usually a way to get around it. An objective basis of comparison, however imperfect, is better than vague generalities. The mere fact that we are talking about new equipment means that we must have some confidence in the future and an anticipation of a continued volume pegged at a reasonable figure. In many cases, reasonable estimates must suffice, but because they may be difficult to determine does not mean those factors do not exist. The prime purpose of this approach is to try to pin down as many objective cost factors as we possibly can evaluate. The final figure we come up with can then be evaluated with the intangibles kept in mind. In this way we narrow down our intangibles rather than throw the whole subject open to the imagination.

Operating-Cost Factors. While many of the afore-mentioned factors are self-explanatory, we should place some special emphasis on the following:

1 When considering labor costs, only measurable savings should be used. Partial reduction of manpower per unit output should not be considered unless one's time can be effectively utilized elsewhere, or if past experience has shown that close attention has been paid to manpower application. In a progressive operation such attention usually results in eventual savings. Some replacements may increase the capacity per worker, but unless this increase is utilized, it should not be considered as a savings.

2 Savings in overhead costs should not be calculated by applying an overhead percentage rate to laborsavings. Only those incremental overhead-cost factors affected by the replacement should be considered. These statements can be explained by considering a replacement that will double production, and the increase can be absorbed by the market. Labor cost in this case has been reduced by one half. However, the annual overhead cost probably has not been reduced; in fact, it may be increased by additional inspection and other service requirements. The replacement may be a desirable one, however, because of the annual laborsavings and the increase of annual dollar profit due to a larger sales volume. Depreciation rates should not be considered in the operating-cost analysis since the effect of capital cost is considered elsewhere in the procedure.

3 Maintenance and repair cost should include the day-to-day equipment maintenance and miscellaneous repair costs, including minor overhauls, necessary to keep the equipment operating properly. Major overhaul cost, which itself constitutes an alternative to the present situation, should not be included in the operating-cost analysis. When considering day-to-day maintenance and repair costs, we are concerned with measuring the difference between the present situation and its proposed alternative, and although this differential will change with time, for practical purposes it is usually sufficient to calculate their present differential.

The proposed replacement may result in either an increase or decrease in the various operating-cost factors. Only the plus or minus differential amounts need to be considered. Their total is called "Annual Gross Savings."

Return on Investment. When talking about equipment replacement we cannot avoid the consideration of return on investment. The advantages of the replacement must do more than just write off its own cost. An enterprise must receive a return on its invested money, whether this be for capital equipment, inventories, or other uses. This concept becomes a little more understandable when we view it as a mortgage taken on the equipment by management, the value of the mortgage being the investment in the replacement. The annual payments on this mortgage are made by the equipment through its annual savings or profit increase. The interest portion of these equal annual payments is the rate of return to be realized from the investment, and is a measure of the earnings potential of the proposed replacement, as compared with the investment of that money in the day-to-day operation of the business. It is a more fundamental measure than annual gross savings, since it is directly related to the investment and the time period involved.

Capital-Recovery Period. Before discussing further the concept of rate of return on investment it will be well to digress for a moment and give the simple arithmetical foundation underlying this method of analysis. As we described previously, a measurement of the capital-recovery period is involved in the first of the three basic tests that a proposed replacement must pass. If this recovery period, which includes the return on investment, reasonably corresponds to the period of anticipated equipment utilization at the volume stated in the analysis, it is logical to assume that the desired return will be achieved. This assumption is no less valid than the usual consideration one must give to the uncertainty of the future. For purposes of demonstrating the basic arithmetic involved we will omit the percentage return on investment for the moment.

Annual gross savings	$ 2,800
Investment	$10,000
Capital recovery period:	
$10,000/$2800 = 3.5 years	

Gross Savings. The second basic test which a replacement should pass is that its gross savings adjusted for wear, after paying for its amortization and return on investment over its service life, must compete with other replacement proposals in their call on capital funds, Table 1.

TABLE 1 SERVICE LIFE—10 YEARS

Annual gross savings	$2800
Less:	
Average annual repair costs over 10 years	500
Adjusted savings	$2300
Less:	
Average annual capital cost ($10,000/10)	$1000
Net average annual savings	$1300

By introducing a return-on-investment requirement of 10 per cent, however, the capital-recovery period is increased to 5 years, and the annual capital cost increased to $1625, with a resulting net saving of $675.

Actually, the return on investment is figured on a diminishing balance in the same manner that a home mortgage is calculated, and is expressed by an annuity formula, shown in most texts on investment or engineering economy

$$A = P\left[\frac{r(1+r)^n}{(1+r)^n - 1}\right]$$

where

A = annual payment provided by savings (annual capital cost)
P = invested capital
r = annual per cent return on investment requirement (compounded at end of year)
n = time period involved, years

This formula can be set up on a nomogram for simple solution, Fig. 1.[1]

Management's Viewpoint

While the determination of a return-on-investment goal is a financial problem, we should have some understanding of its meaning. The rate of return requirement ties in with the total operation of the business and a specific goal for an enterprise depends on many factors, including the type of business, the rate of its return on existing investment, the availability of funds, and, particularly, the viewpoint of management in regard to the employment of capital. Because the value of the return rate can affect a replacement decision, this matter of management viewpoint is crucial.

Extremes must be avoided, for they are likely to distort judgment. If it is to be a reliable guide, the return rate will have to reflect factors related to the use of the capital as well as its acquisition. Further, a replacement opportunity under consideration will not be expected to return any more, or for that matter, any less, than the going or desired rate of return applying to the employment of capital in other phases of the business. Of course the final decision on a capital expenditure should take the maintenance of the cash position into account; and of course safety factors are not necessarily inappropriate if they are introduced into the analysis at points where they can be recognized for what they are. But the return rate itself must be kept pure and sound.

Perhaps the importance of looking at the return rate in this way can be better understood when it is realized that we are measuring a given proposal against all alternatives, including other possible replacement proposals. The rate of return goal affects all proposals equally and adjusts for the time value of money. It does not attempt to dictate which replacement has first call on available capital funds.

[1] "How to Figure Equipment Replacement," by P. A. Scheuble, Jr., *Harvard Business Review*, Boston, Mass., September–October, 1955.

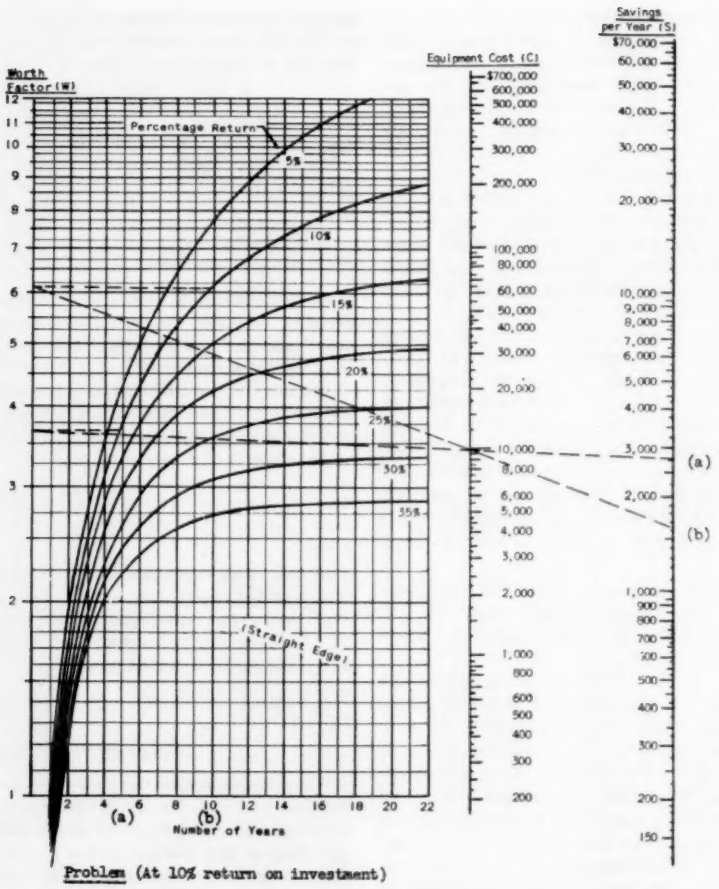


FIG. 1 NOMOGRAM RELATING SAVINGS PER YEAR, EQUIPMENT COST, AND NUMBER OF YEARS TO PAY OFF AN INVESTMENT PLUS A GIVEN PERCENTAGE RETURN ON INVESTMENT

SERVICE LIFE

Thus far we have mentioned two types of time periods: The capital-return period and service-life period. It is thought the concept of capital-return period is comparatively easy to understand; however, service life will require a little closer analysis to get a better understanding of its meaning. The belief that an asset has an inherent service life independent of the duration of its specific application is misleading. Equipment obsolescence can have a marked effect on service life, regardless of how excellent its condition. In other cases, however, the effect of obsolescence may be small in comparison to that of wear. In most problems, however, both obsolescence and wear are involved, but their effects must be evaluated separately since they usually do not occur simultaneously.

Obsolescence Versus Wear. Obsolescence cost is determined by a comparison with a proposed replacement that is superior in operation, while wear, on the other hand, occurs gradually through usage. For practical purposes, the savings afforded by the proposed replacement would hold throughout the service life if wear did not take place. However, wear does occur, and the dollar amount must be predicted quantitatively in order to adjust for its effects on potential operating savings.

Wear is a physical phenomenon, but it can be reasonably measured by repair cost. Most manufacturing and related equipment consists essentially of a structure that supports, and also may be in contact with moving parts. Since these parts wear, they may be replaced to maintain the original operating characteristics of the equipment. We also might have structural distortions, corrosion, and wear of bearing supports. The incidence and amount of repair costs are difficult to determine, but they are usually low in early life and increase with age. Also, they vary with the severity of use. Since their exact growth pattern is indeterminate, we may for the sake of convenience assume that their annual cumulative average for a given period increases in proportion to equipment age, or hours of use, without introducing any critical errors, Fig. 2.

Establishing Service Life. Thus far we have not come up with the answer as to when obsolescence or wear establishes the service life. Of course the problem is relatively simple for a piece of equipment peculiar to a contract job that will run for a specified

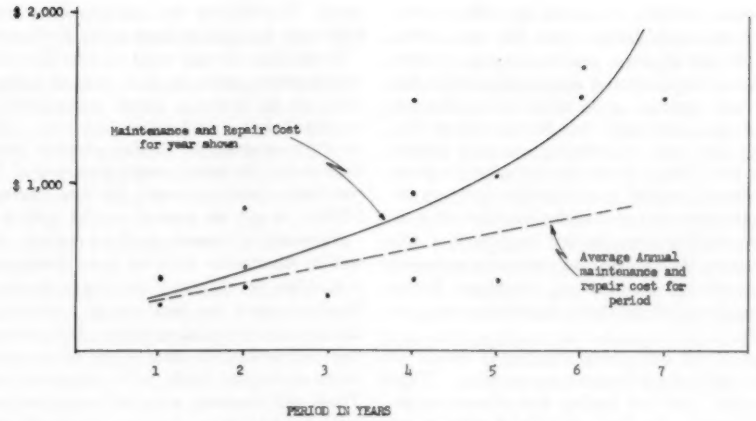


FIG. 2 MAINTENANCE AND REPAIR COST

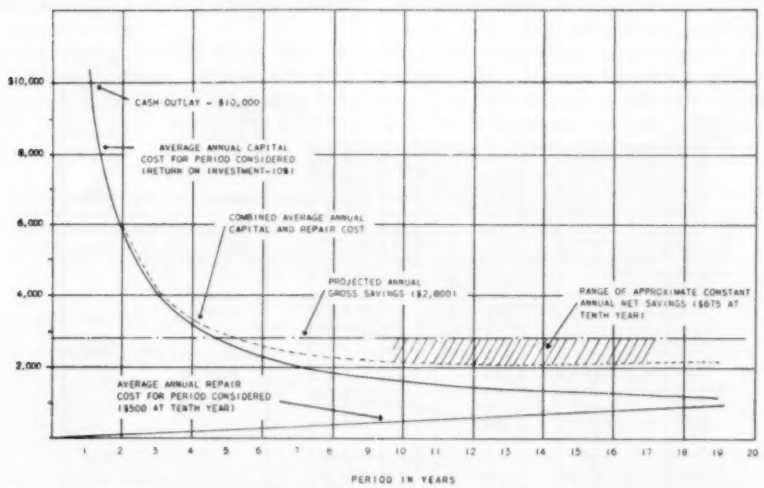


FIG. 3 COMBINED EFFECT OF AVERAGE ANNUAL CAPITAL COST AND AVERAGE ANNUAL REPAIR COST ON ANNUAL GROSS SAVINGS

period. Generally, however, the determination is not so simple because both wear and obsolescence are involved. Perhaps this question can be resolved if we pose a still more basic question: When can we anticipate making a comparatively large capital outlay for either overhaul or replacement?

In some few cases we may be able to make such a prediction if we know that certain trends in equipment design will make superior equipment available in the foreseeable future. In more cases, however, we can reasonably predict that under given operating conditions certain classes of equipment will require a major overhaul after a definite period of use, and that the comparatively large expenditure required at that time may make it desirable to consider the purchase of superior equipment. Thus the time of major overhaul represents a point of potential obsolescence as well as a major stage of physical deterioration.

Point of Diminishing Returns. This reasoning can be applied to a great number of types of equipment in a company that has a progressive equipment policy and is trying to get the best in equipment performance per dollar it spends. It is not arbitrary in the light of the fact that repair-cost patterns have a direct effect on the decision to overhaul, and the fact that as repair costs increase the need for overhaul or replacement becomes more urgent. A graphic representation of average annual capital costs and repair costs will show that their combined total will reach a minimum point, after which it becomes less economical to operate a piece of equipment unless the repair cost is reduced by overhaul or replacement, Fig. 3.

The net savings in this example remained fairly constant over a wide span, and a service life selected in this range would for practical purposes give savings of the same magnitude. This result reflects realistically the problem usually presented when deciding to overhaul or buy. The range of error is acceptable on practical grounds. Seldom can we decide exactly when to take action, but usually we can make a decision within a reasonable period without adverse effect. Such a decision can be influenced by the present or anticipated availability of equipment, the price level of equipment, imminent improvements in design, as well as our decisions in regard to the priority of a replacement in relation to capital funds.

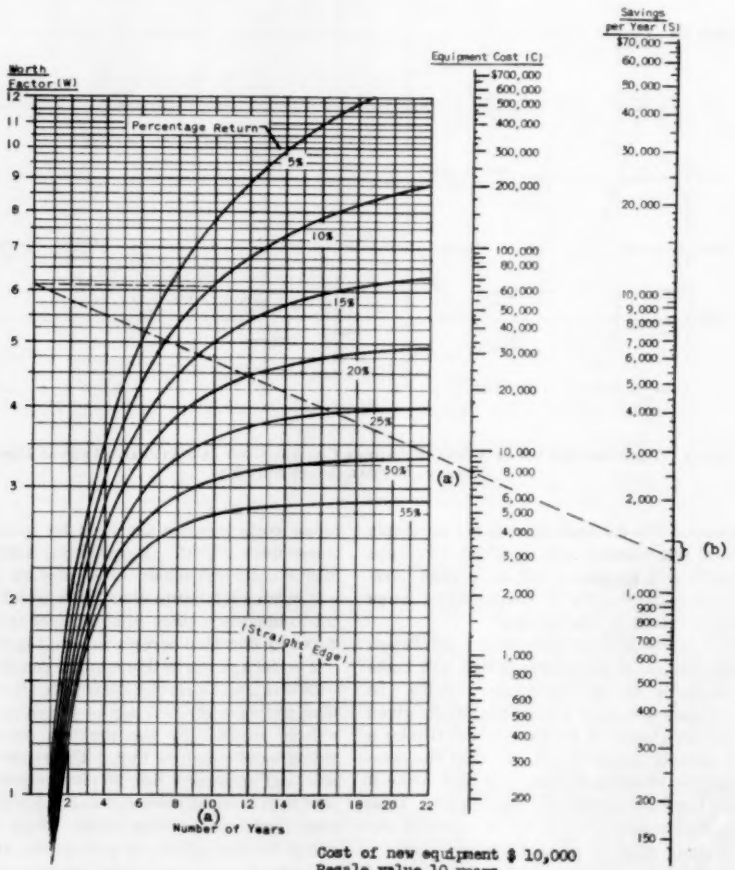
Salvage Value. An equipment replacement decision is usually concerned with the net outlay of cash involved, and the resale or salvage value of the replaced equipment reduces the cash out-

lay. This raises a question whether we should also consider the future salvage value of the replacement, since this will reduce annual capital costs. We can of course make a guess as to what the future salvage value of the proposed replacement might be, but this requires not only guesses as to when the equipment might be available for resale, but also what future market conditions might be. We are only introducing guessing factors against which the real cash outlay of the present must be measured. If future value has a bearing on a selection between alternative pieces of equipment, it is far better to think of it in terms of comparative salability than to try to give it a numerical measure. After all, the typical manufacturing company is not in the used-machinery business so, therefore, future values should not generally be of any great significance in a replacement decision.

Both present values of old equipment and future values of new equipment should not be used in the same problem. That would be double counting, if we are dealing with a series of replacements over many years. It's just a matter of whether we want to count the future salvage value now in figuring capital cost or to wait until that asset is sold and credit the new replacement. Considering the indeterminate nature of salvage or resale value the cash-in-hand point of view is a little more certain.

Sometimes we may want to take the future salvage value into consideration, such as in a case of rather short-term installation, or to make a rough comparison between rental and buying (before consideration of taxes). This can be worked out on the nomogram for the first problem (assume $2000 future salvage value) by using a net investment of $8000 and adjusting the resultant answer upward for the interest on $2000 (10% × $2000), to get an annual capital cost of about $1500, Fig. 4.

Summary of Reasons for Replacement. At this point it will be well to summarize what we have discussed, since we are now in a position to approach making a decision on a replacement. Having passed the first test of a reasonable utilization during the period of capital recovery, the proposed replacement must then compete with other worth-while replacements in their demand on capital funds, by a comparison of net annual savings. Those replacements with the largest net savings should receive prior consideration; however, it should be emphasized that the net-savings figure by itself is not the criterion which determines the final decision, but it is only a valuable gage. Sometimes



Cost of new equipment $ 10,000
Resale value 10 years from now 2,000
(a) Net $ 8,000

(b) Adjust for time value of resale value $ 2,000 x 10% equals $ 200

FIG. 4 NOMOGRAM RELATING SAVINGS PER YEAR, EQUIPMENT COST, AND NUMBER OF YEARS TO PAY OFF AN INVESTMENT PLUS A GIVEN PERCENTAGE RETURN ON INVESTMENT

there are intangibles which we must balance against the net annual savings.

There is another point to remember. In many cases, equipment decisions can be classified as those which are of a mandatory nature; that is, we must make up our minds now or within a reasonable period, such as in the case of a major overhaul or buying new. In such cases the capital-recovery period is an academic consideration. Another type of decision usually involves a replacement that is more economic in nature, since production may not be jeopardized, although we can lose money by not making that decision.

In the first case the decision may revolve around a comparison of the annual capital costs of overhaul versus buy, with, perhaps, operating savings entering into the comparison. In the second instance the decision is usually based upon the operating savings involved in comparison to the annual capital cost. This difference must be kept in mind in comparing the net-savings figures. Obviously, if the income loss due to production shutdown were taken into consideration in the mandatory-type analysis, the net-savings figure would generally beat out the replacement being considered only for its cost savings.

Also entering into the replacement decision will be the comparative risk involved. Generally, those replacements with short capital-recovery periods will show better net savings, thus simplifying the problem of selection. Sometimes, however, equal net savings are available from other projects with longer recovery periods. While no rigid rule can be established in choosing between such alternatives, the replacement with the shortest recovery period may be more desirable from the standpoint of less time risk. Again, however, there is no substitute for good judgment; there are plenty of factors that can make the opposite decision the wise one.

The example illustrated in Figs. 5, 6, and 7 is shown more for the points involved, rather than to cover the multitude of possible replacement situations. In calculating the three-year return period in the foregoing problem the effect of repair cost was not considered, since over this length of time it would not be significant. When long return periods are involved the annual gross savings should be reduced by the average predicted repair cost for the calculated period, and the recovery period resolved. This is a method of successive approximations, but is sufficiently accurate.

A Planned Program

This approach to replacement analyses requires that a planned program be in effect, since a comparison between replacements cannot be made unless all possibilities are surveyed prior to selection, It is believed that this is only logical and practical, since it is the only way to avoid the helter-skelter expenditure of money, and perhaps at a later date some worth-while project

PRESENT EQUIPMENT

Manufacturer CINCINATTI — BICKFORD
Type and Size 12-30 DUPLEX MILL and DRILL
Machine No. VERTICAL DRILL PRESS
Year Built 1944 & 1930

NEW EQUIPMENT

Manufacturer CINCINATTI
Type and Size SINGLE RAM BROACH
Model Description No. S-54
Estimate by JAB Date 12-30-57

ANNUAL AVERAGE OPERATING COMPARISON OF PROPOSED ALTERNATIVE vs. PRESENT EQUIPMENT

COST ITEMS	Savings of proposed over present	Additional Costs over present
1. Direct Labor	$ 6480	$
2. Indirect Labor	300	
3. Fringe Benefits (Use 15% of 1 and 2)	1020	
4. Maintenance and Repair (exclude major overhaul)	250	
5. Tools (Production, durable and perishable)		
6. Scrap and Rework	50	
7. Down Time		
8. Power Costs		
9. Subcontract Cost		
10. Property Taxes and Insurance		600
11. Other items such as: Income increase due to increased capacity (if utilized), material cost, cost of space (if significant) Specify		
Total	$ 8100 (A)	$ 600 (B)
12. Annual Gross Savings of proposed over present (A minus B)		$ 7500

CAPITAL COST ANALYSIS

PRESENT EQUIPMENT		NEW EQUIPMENT (or alternative of overhaul)	
13. Resale, Salvage or Conversion value	$ 1050	17. Anticipated Service Life 10 years	
14. Remaining useful life for this application before necessary major overhaul	1 years.	18. Cost of Equipment (include engineering)	$ 15350
15. Cost of major overhaul	$ 6000	19. Installation and rearrangement cost	1500
16. Useful life after overhaul	5 years.	20. Cost of additional items needed to service equipment	
NOTE: Attach separate Machinery and Equipment Sheet for alternative of overhaul		21. "Going-in" Costs	
		22. Other expensed items	
		23. Total cost of replacement	16850
		24. Subtract item (13)	1050
		25. Net cost of replacement or overhaul	$ 15850

SUMMARY (See Instructions)

NUMBER OF YEARS TO RETURN INVESTMENT		AVERAGE NET SAVINGS DURING SERVICE LIFE	
26. Required % return on investment	20 %	29. Annual gross savings - From (12) above	$ 7500
		29. Less average anticipated repair costs	384
27. Number of years to return investment at % required return (Use Nomogram)	3 years	30. Adjusted gross savings (28) minus (29)	$ 7116
		31. Annual savings required to give desired return for period of anticipated service life in (17) above (Use Nomogram)	3750
		32. Annual average Net savings during anticipated service life (30 minus (31))	$ 3366

Comments: Three year return period falls within this UTILIZATION Period. Also overhaul is less economical than buy.

NOTE: NET SAVINGS FOR overhaul is (+) $3750

Approved by John S. Johnson

FIG. 5 MACHINERY AND EQUIPMENT ANALYSIS

PRESENT EQUIPMENT

Manufacturer _____
Type and Size _____
Machine No. _____
Year Built _____

NEW EQUIPMENT

Manufacturer OVERHAUL OF DUPLEX MILL
Type and Size AMP VERTICAL DRILL PRESS
Model Description _____
Estimate by JR Date 4-30-54

ANNUAL AVERAGE OPERATING COMPARISON OF PROPOSED ALTERNATIVE vs. PRESENT EQUIPMENT

COST ITEMS	Savings of proposed over present	Additional Costs over present
1. Direct Labor	$ 1000	$
2. Indirect Labor		
3. Fringe Benefits (Use 15% of 1 and 2)	150	
4. Maintenance and Repair (exclude major overhaul)	250	
5. Tools (Production, durable and perishable)		
6. Scrap and Rework		
7. Down Time		
8. Power Costs		
9. Subcontract Cost		
10. Property Taxes and Insurance		
11. Other items such as: Income increase due to increased capacity (if utilized), material cost, cost of space (if significant) Specify		
Total	$ 1400 (A)	$ (B)
12. Annual Gross Savings of proposed over present (A minus B)		$ 1400

CAPITAL COST ANALYSIS

PRESENT EQUIPMENT	NEW EQUIPMENT (or alternative of overhaul)
13. Resale, Salvage or Conversion value $ _____	17. Anticipated Service Life 5 years
14. Remaining useful life for this application before necessary major overhaul _____ years.	18. Cost of Equipment (include engineering) $ 6000
15. Cost of major overhaul $ 6000	19. Installation and rearrangement cost
16. Useful life after overhaul 5 years.	20. Cost of additional items needed to service equipment
NOTE: Attach separate Machinery and Equipment Sheet for alternative of overhaul	21. "Going-in" Costs
	22. Other expensed items
	23. Total cost of replacement
	24. Subtract item (13)
	25. Net cost of replacement or overhaul $ 6000

SUMMARY (See Instructions)

NUMBER OF YEARS TO RETURN INVESTMENT	AVERAGE NET SAVINGS DURING SERVICE LIFE
26. Required % return on investment _____ %	28. Annual gross savings - From (12) above $ 1400
27. Number of years to return investment at % required return (Use Nomogram) _____ years	29. Less average anticipated repair costs 150
	30. Adjusted gross savings (28) minus (29) $ 1250
	31. Annual savings required to give desired return for period of anticipated service life in (17) above (Use Nomogram) 2000
	32. Annual average Net savings during anticipated service life (30 minus (31)) $ (-) 750

Comments

Approved by _____

FIG. 6 MACHINERY AND EQUIPMENT ANALYSIS

may have to go by the board because we've used up our capital budget for that year. Also, it is not a good way to create a spirit of confidence in our proposals. In fact, an annual review of possible replacements is desirable. Management can then evaluate the savings potential and plan the allocation of funds or acquire the necessary capital.

Repair Costs. We should briefly consider some mathematical effects of treating repair costs in this method, and the use of annual costs:

1 We do not consider the time value of money relationship on repair costs. The effect is to inflate repair cost. However, since the pattern is indeterminate, and generally, actual repair costs are usually understated on records, the practical effect is negligible in comparison to other factors that will dictate replacement.

2 The annual gross savings should be based on an average value of utilization. Theoretically, if year-to-year values are not equal, an adjustment must be made for the time value of money. In effect, savings have a larger present worth in the earlier years, so in the case of a gradually declining volume of production, the annual gross-savings figure is only slightly understated by an unweighted average. Where there are extreme cases, a present worth[8] basis of analysis is desirable.

[8] "Principles of Engineering Economy," by E. L. Grant, The Ronald Press Company, New York, N. Y., third edition, 1950, chapters 7 and 8.

Effect of Taxes. While a replacement decision should not be primarily concerned with taxes, we cannot completely ignore this item. Sometimes, however, too much emphasis is put on the tax aspect. It's like medical expenses: It's nice to know you have a deduction after the necessity arises, but one usually has more money left by not getting sick in the first place. It is important to have some understanding of this effect.

The fact that replacement is designed to increase profits means that taxes will be higher as a result. However, applying a flat tax-percentage effect on gross savings neglects the effect of other factors in the profit-and-loss statement. It seems only sensible to consider the tax effect through the aggregate profit-and-loss statement in relation to replacements in general. After all, our purpose in the procedure described here is to compare the desirability of alternative replacements, and the tax on increased income would hardly change their relationship. Should it be desired to consider the tax effect in the calculation, the percentage-return-on-investment figure would be adjusted upward to include the nominal effect of taxes in relation to investment for the enterprise.

Considerations of Accounting. The value of the capital expenditure in a replacement analysis will include expenses not normally classified as capital items, which in the usual accounting treatment would be charged directly against profit and loss and thus would reduce taxes. The accounting treatment, of course, makes this distinction between investment costs, which are depreciated over time, and operating costs which are expended in the current year, whereas a replacement analysis seeks to measure the cash outlay of a replacement over the whole period of its utilization. Here accounting is simply serving a different purpose; it must maintain stability in the determination of profits from year to year, and it must take account of tax requirements.

Depreciation. There may be more point in not ignoring the effects of depreciation on income taxes, for these will vary between contemplated replacements. The higher depreciation charges made possible by the new capital asset will tend to reduce taxes and will thus reduce the total net cash outlay of the company; and the bigger the asset, the greater the effect. Further, when the asset is removed from the capital-asset account at the end of its useful life to the company, the undepreciated amount will tend to reduce taxes; though the reverse possibility also exists that the salvage or resale value received may exceed book value and thus tend to increase taxes.

The practical fact is, however, that such tax effects may be indeterminate for a particular problem. Thus:

1 The Internal Revenue Code of 1954 provides for a company's taking larger depreciation allowances (up to 200 per cent of straight-line depreciation in the first year of acquisition of an asset), which can act to defer taxes as well as to accumulate

reserves that are more commensurate timewise with the real depreciation of the equipment. Undoubtedly, the full advantage of these liberalized depreciation provisions can be realized by companies that follow a vigorous replacement policy and have asset-accounting procedures that can take advantage of this law. In general, however, the possible advantages cannot be reduced to clear-cut rules.

2 Again, the net salvage value of one replaced machine tool may be offset by the write-off of others, and hence no net capital gain is involved. It would, therefore, be difficult to set up a general rule regarding salvage value to guide the operating organization in analyzing this factor.

While any such tax effect, if significant and measurable, should of course be considered, it should not be injected into the replacement study where it is only likely to confuse the analysis. It is usually sounder to make the analysis primarily on the basis of good engineering and business judgment, with the tax effects brought to bear subsequently.

Making the Replacement Pay Off

To make equipment replacement pay off we must have a coherent plan, if we are to get the most for our investment. Such a plan should include:

1 An objective study of proposals, and supporting them by records or reasonable estimates. Such an analysis also includes a periodic check of equipment condition and operating costs which may signal deterioration. Good working records are required, but they need not be expensive since they usually can serve multiple purposes. A good example is the maintenance record that can be used both for historical information and as a working record to schedule maintenance.

2 An economic analysis that will tie together the engineering, cost, and financial factors involved. A priority and timing of replacements should be established.

3 A two-way communication among, and between, all levels in the organization, so that plans can be correlated with future company planning and capital availability.

4 A broad management viewpoint in regard to the use of funds for replacement purposes, the setting up of adequate reserves for replacement, and an appreciation of its competitive position in regard to the efficiency of its plant.

Date 12-30-54

(Note: Show details of cost items on separate sheets)

SALES VOLUME ANALYSIS

Express volume in terms of units (or hours) processed over equipment considered. Where more than one item or product is involved (Product Mix) use equivalent units (or hours). Units can be pieces, pounds, etc.

33. Volume for past years (Fiscal) (Calendar) (19[5]) 32,000 (Units) ~~(Hours)~~
(19[5]) 46,000 (Units) ~~(Hours)~~
(19[5]) 60,000 (Units) ~~(Hours)~~
34. Average yearly volume projected for this analysis 69,000 . For next 5 years.
35. Replacement will be utilized approximately 20 % on a 2 shift basis.

Volume projection is conservative because of expanding market. Other items can also be run on this equipment.

DIRECT LABOR SAVINGS

Measure savings in terms of operators wages at present contract rates. (Consider overtime and shift premiums). Where units are involved use standard labor per unit.

36. Present cost per (Unit) ~~(Hour)~~ $ 1.25
37. Proposed cost per (Unit) ~~(Hour)~~ 1.142
38. Savings per (Unit) (Hour) $ .108
39. Will labor standards be changed? (Yes)
Other calculations (attach sheet)

Saving per year for next 5 years is $ 6480 (Item 1)
NOTE: (38) multiplied by (34)

INDIRECT LABOR SAVINGS (YEARLY)

Itemize cost of following, if changed:

	PRESENT	PROPOSED	DIFFERENCE
Supervisory	$	$	$
Inspection			300
Material handling			
Other indirect			
TOTALS	$	$	$ 300 Item (2)

GOING-IN COST

"Going-in" costs are above normal costs in connection with getting equipment or process running. (Excess labor over standard, excess scrap losses, excess down time, excess tool tryout, reworking of equipment or tooling etc.).

Explanation and calculations:

Fig. 7 Summary Work Sheet

5 An after-the-fact review of proposals installed. Sometimes our dream boat is rocked, but the review is important since it gives us improved judgment and may avoid future mistakes.

6 A program that will provide proper equipment maintenance, and constant supervisory effort to instruct and prevent mishandling of equipment. Needless to say, this factor is of great importance.

We have discussed the economics of equipment replacement. However, it may be recognized clearly, that what we have discussed is only one facet of the day-to-day problem of increasing the manufacturing efficiency of a company. In the background must be the continuing review of the things we are doing and how we can do them in a better way, increase quality, reduce fatigue, make operations safer, and reduce cost. The fact remains that such ideas must be evaluated in terms of understandable ground rules, so that they can be compared for decision and their acceptance or rejection be explained. It is believed that if we follow some of these principles, we can show that equipment replacement can and does pay off.

Applying Linear Programming to Inventory Planning in a Seasonal Market A Case Study

By D. W. MOFFETT,[1] PITTSBURGH, PA.

This paper outlines a method for planning the most economical production and inventory program for maintaining stable employment levels in a seasonal market.

Introduction

CYCLIC variations of demand occur in many industries and for many reasons. A great deal of study and research work has been done to try to predict seasonal-market fluctuations so that manufacturers of goods might plan their production accordingly. These efforts have met with considerable success. Many companies can now forecast seasonal demand peaks with a reasonable degree of accuracy.

Cyclic variations are not limited to items produced for a seasonal market, however. In some plants, the demand for parts for an assembly may be cyclically variable because the assembly is produced only in certain seasons. In the automobile industry, for example, the annual model change causes a slump in the demand for parts similar to a seasonal-market variation.

Regardless of the cause of the variation, cyclical fluctuations in demand complicate the planning problem. Production facilities which are capable of meeting seasonal peaks must operate at fractional capacities in periods of minimum demand. On the other hand, stabilizing production at the average demand rate results in accumulation of inventory during slack seasons. Either of these alternatives is costly.

In recent years, manufacturers are finding it more and more difficult to plan flexibility in production capacity. Current labor conditions are increasing the penalties for variations in the size of the labor force so that management is finding it necessary to plan for stabilized levels of employment even though the demand for products may be fluctuating.

When production is planned to provide steady employment, inventory accumulation becomes necessary to absorb excess capacity in slack seasons, and later to provide goods to satisfy peak demands. The cost of carrying this accumulation is an important factor in planning a stabilized production program.

The advent of "linear programming" has provided management with a new tool which is valuable in dealing with planning problems of this type.

Linear programming is a mathematical technique for determining an optimum program of action when the situation is affected by many variables. In several instances, linear programming has been applied to production-planning problems to determine a most profitable, least costly, or most productive program.

This paper describes an application of linear programming to the problem of providing stable levels of employment in a plant producing parts for an assembly line which operates at a seasonally variable rate. Linear programming makes it possible to formulate a plan for maintaining stable production levels with minimum cost.

Conditions of the Problem

This case example concerns a screw-machine department which produces parts for several finished products. Demand for the finished products is subject to seasonal variations because of the Christmas trade.

About 210 different screw-machine parts are required for all of the products. Most of these are typical screw-machine items which can be procured from any screw-machine jobbing shop. About 35 of the parts, however, are of some special nature or involve quality considerations that make it impractical to entrust them to a vendor.

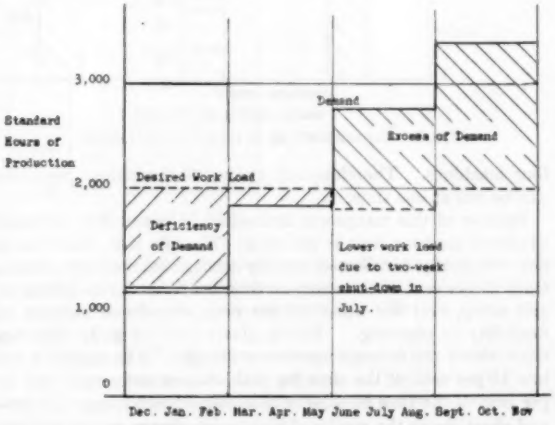


Fig. 1 Comparison of Demand and Desired Work Load During One Seasonal Cycle

The seasonal variation of the demand for parts is shown in Fig. 1. The solid line on this graph shows the number of standard hours of production required in each quarter to provide screw-machine parts according to the assembly schedule. The broken line shows the number of standard hours of production that must be assigned in each quarter to provide the desired level of employment. The shaded areas between the lines show the deficiency of demand in the first two quarters of the year and the excess of demand in the last two. This graph makes it apparent that, if the desired work load is to be maintained, some of the demand of the third and fourth quarter must be satisfied by production in the first and second quarters. This will cause an inventory accumulation through the middle of the year. The remainder of the excess demand must be satisfied by purchasing. Since these purchases can be made as the parts are needed, only normal inventory levels are required.

Sixteen screw machines of three sizes are available in the shop. The operating crew consists of three men, each of whom runs

[1] Senior Consultant, Methods Engineering Council.

Contributed by the Management Division and presented at the Engineering Management Conference, St. Louis, Mo., March 14–15, 1956, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, December 6, 1955. Paper No. 56—MGT-1.

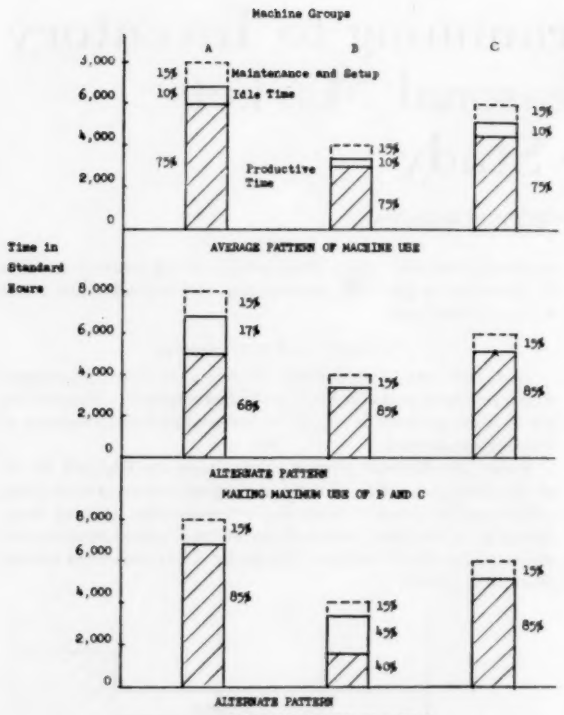


FIG. 2 FLEXIBILITY OF MACHINE UTILIZATION

four machines. Therefore any twelve of the sixteen machines can be run at one time.

Because of this manpower limitation, 25 per cent of the total available machine time is idle time. The idle time, however, is not necessarily distributed equally among the machine groups. Only 15 per cent of the machine time is needed for maintenance and setup, and the other 10 per cent provides a measure of flexibility in planning. This is illustrated in Fig. 2. The top chart shows the average machine utilization. The machines are idle 15 per cent of the time for maintenance and setup, and 10 per cent of the time because of manpower limitations. The second chart shows the results of making maximum use of machine groups B and C. It is necessary to have 32 per cent idle time on group A. The bottom chart shows the pattern when maximum use is made of groups A and C. Group B is idle 60 per cent of the time.

Linear programming makes it possible to consider this flexibility in planning the production schedule of the department. The calculated program applies the available manpower to the machines in the way that best serves the stated purpose.

Each part made in this shop is tooled for running on only one size of screw machine. This simplifies the problem somewhat. It is possible, however, to use this same linear programming method to consider parts that can be run on alternate machines under some circumstances. This will be shown later.

ADVANTAGES OF LINEAR PROGRAMMING IN THIS APPLICATION

In this application, linear programming is used to determine when, and in what quantities, each part should be made or purchased to provide a stabilized work load for the labor force at minimum cost. With a stabilized work load, the expense of hiring and laying off personnel to accommodate market fluctuations is eliminated. In addition, the risk of losing skilled employees to other companies and the consequent expense of training replacements is reduced. A stable work load also promotes better employee relations and public relations, although these values are more difficult to measure in dollars.

Using linear programming to plan production for a stable work load has several advantages as follows:

1 Linear programming provides a least-cost program. Management is assured that the calculated plan is the most economical possible under the circumstances.

2 Linear programming evaluates the total cost of leveling the work load. With this information, management can compare the costs of maintaining various constant levels of production and employment with the costs of layoffs, rehiring, training, unemployment compensation, and the like. This information provides management with a basis for evaluating the costs of programs of guaranteed annual wage.

3 The calculated program is specific. Machines, parts, quantities, and time are all specified.

4 By-products of the calculation provide management with vital information on which to base decisions. The desirability of adding production capacity is typical of the factors about which factual, quantitative information is developed.

5 Calculation of the program is simple and rapid. Alternative plans can be worked out on several different forecasts to show how management can "hedge" against forecast errors, and plan for minimum inventory risk in rising and falling markets.

The application to be described demonstrates how these advantages are obtained.

DEFINITION OF THE PROBLEM

The problem of planning leveled production in this shop can be defined as follows:

The requirements for a part in any one quarter may be satisfied in three ways:

(*a*) Make the part in the quarter in which it is needed.
(*b*) Make the part in a previous quarter and carry it in inventory until needed.
(*c*) Purchase the part when it is needed.

What program of production and procurement will maintain the desired level of work load at least cost?

For each of the three ways of supplying a part, the cost is different. The problem is to plan the method of providing each part so that the seasonal demand is satisfied and the total cost is a minimum.

DESCRIPTION OF A SMALL MODEL PROBLEM

For convenience of illustration, consider the small model problem shown in Table 1. The upper table shows the cost and demand data for seven parts. The column headed "Machine Group" shows which of three machine sizes—A, B, or C—is normally used for making the part. The next column shows alternate routings for parts II and IV. Part II can be made on group A or B, and part IV can be made on group B or C. The next column shows the penalty per standard hour that is incurred by making the part by the alternate routing. This value is simply the difference between the costs per standard hour of operation of the two routings.

The next column shows the cost of carrying one hour's production in inventory for one quarter. For example, the cost of carrying one hour's production of item I for three months is two dollars. The cost of carrying inventory has been assumed to be directly proportional to the value of the item, although adjustments can be made for unusually high risks of deterioration or obsolescence without changing the method of calculation.

TABLE 1 DATA FOR MODEL PROBLEM

DEMAND FOR PARTS

Part	Mach. Group	Alt. Mach. Group	Pen. /Hr.	Cost of Carrying 1 Hour's Prod. for 1 Quarter	Mfg. Profit Advantage /Hour	Hrs. Demand by Quarters			
						1	2	3	4
I	A			2.00	must be made	200	400	600	700
II	A	B	1.20	2.00	4.00	100	100	300	200
III	B			.50	2.00	200	200	200	400
IV	B	C	1.00	1.50	1.50	300	600	900	1050
V	C			1.00	1.00	50	50	150	100
VI	C			1.50	-1.00	300	300	300	600
VII	C			2.00	-3.00	100	200	300	350
				Total Demand		1250	1850	2750	3400
				Desired Work Load		2000	2000	1800	2000
				Net Adjustment Required		-750	-150	-950	1400

MACHINE CAPACITY AVAILABLE

Mach. Hours Available	Quarter				Total for Yr.
	1	2	3	4	
A	1000	1000	900	1000	3900
B	500	500	450	500	1950
C	1000	1000	900	1000	3900
Total in each quarter	2500	2500	2250	2500	9750
Man Hours Available	2000	2000	1800	2000	7800
Idle Time Necessary	500	500	450	500	1950

The column headed "Mfg. Profit Advantage/Hour" shows the penalty incurred when an amount of the part equivalent to 1 standard hour's production is purchased instead of made. Notice that part I is marked "must be made." Also, the values for parts VI and VII are negative, indicating that they can be purchased more cheaply than they can be made.

The four columns to the right of the table show the demand for each part in each quarter expressed in standard hours of machine time. For example, the demand for part I requires 200 hr of machine time in the first quarter, 400 hr in the second quarter, 600 in the third, and 700 in the fourth.

At the bottom of the four columns are shown the total demands for machine hours in each of the four quarters. Below these totals are shown the desired work loads for each quarter; and below these are the deficiencies and excesses.

The upper part of Table 1 gives a complete summary of the conditions. The demand for each item in each quarter, and the relative cost of any of the three ways of supplying the demand are all available from the table.

The lower table shows the limitations imposed by the availability of machine time in each quarter. Machine groups A and C each have 1000 hr available in the first, second, and fourth quarters. Only 900 hr are available in the third quarter because of the summer shutdown. Machine group B has just one half as many hours available as either A or C. This provides a total of 2500 machine hr in the first, second, and fourth quarters—slightly less in the third.

The line headed "Man Hours Available" shows the amounts of machine time that the manpower limitations will allow to be used. The next line shows the idle time that will be necessary as a result. In the first, second, and fourth quarters, 500 hr of idle time must be distributed among the three groups. In the third quarter, 450 hr must be idle.

This is all of the data necessary to apply the Modi method of linear programming to this problem. The Modi method considers simultaneously all of the restrictions and costs and arrives at the most economical program.

The most economical program is shown in Table 2. The number of hours of production scheduled in each quarter is shown for each part. The number of hours purchased is shown in the last column.

For example, see part I at the top of the table. The first-quarter requirement and 200 hr of the second-quarter requirement are made in the first quarter. The remaining 200 hr of second-quarter requirement and 200 hr of third-quarter requirement are made in the second quarter. In the third quarter, the remainder of the third-quarter requirement is made, and the entire fourth-quarter requirement of 700 hr is made in the fourth quarter.

This program requires that 200-hr production of part I be carried in inventory from the first quarter to the second, and 200 hr from the second quarter to the third.

The complete production program for each part is laid out in this solution. Also the purchases of the fourth-quarter requirement of part VI and the second, third, and fourth-quarter requirement of part VII are shown in the last column. The bottom line of the table shows how the required amount of idle time is distributed among the machine groups. In the first and second quarters, all of the idle time is assigned to group A. In the

TABLE 2 MOST ECONOMICAL PROGRAM

PRODUCTION PROGRAM

Part	Qtr.	Req't.	First Quarter			Second Quarter			Third Quarter			Fourth Quarter			Purch.
			M-A	M-B	M-C	M-A	M-B	M-C	M-A	M-B	M-C	M-A	M-B	M-C	
I	1	200	200												
	2	400	200			200									
	3	600				200			400						
	4	700										700			
II	1	100	100												
	2	100				100									
	3	300							300						
	4	200										200			
III	1	200		200											
	2	200					200								
	3	200		200											
	4	400		100			300								
IV	1	300			300										
	2	600						600							
	3	900								450	450				
	4	1050											500	550	
V	1	50			50										
	2	50			50										
	3	150			150										
	4	100			50									50	
VI	1	300			300										
	2	300						300							
	3	300						100			200				
	4	600													600
VII	1	100			100										
	2	200													200
	3	300													300
	4	350													350
Idle Time			500			500			200		250	100		400	

Note: M = Machine

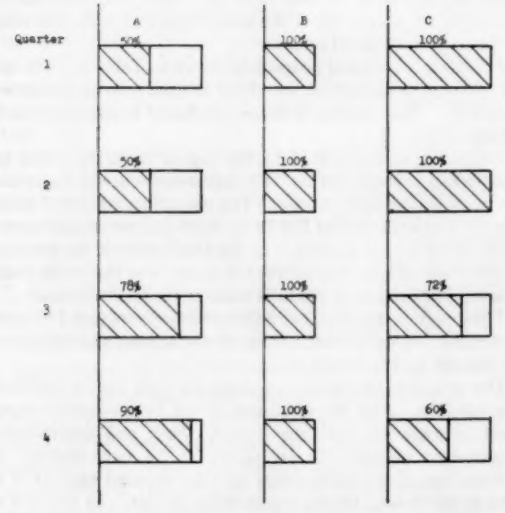
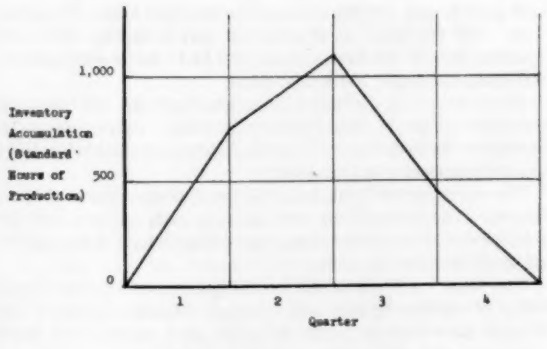


FIG. 3 MACHINE UTILIZATION IN MOST ECONOMICAL PROGRAM



Cost of carrying

Inventory from	1st to 2nd quarter	$800
	2nd to 3rd quarter	1050
	3rd to 4th quarter	250
	Total	$2100

FIG. 4 INVENTORY ACCUMULATION RESULTING FROM MOST ECONOMICAL PROGRAM

TABLE 3 PRODUCTION PROGRAM WITHOUT LEVELING

PRODUCTION PROGRAM

Part	Qtr.	Req't.	First Quarter			Second Quarter			Third Quarter			Fourth Quarter			Purch.
			M-A	M-B	M-C	M-A	M-B	M-C	M-A	M-B	M-C	M-A	M-B	M-C	
I	1	200	200												
	2	400				400									
	3	600							600						
	4	700										700			
II	1	100	100												
	2	100				100									
	3	300							300						
	4	200										200			
III	1	200		200											
	2	200					200								
	3	200								200					
	4	400											400		
IV	1	300		300											
	2	600					300	300							
	3	900								250	650				
	4	1050											100	950	
V	1	50			50										
	2	50						50							
	3	150									150				
	4	100												50	50
VI	1	300			300										
	2	300						300							
	3	300									100				200
	4	600													600
VII	1	100			100										
	2	200						200							
	3	300													300
	4	350													350
Idle Time			700		550	500		250							
Layoff Hrs.			750			250									

Note: M = Machine

second and third quarter, it is divided between groups A and C.

The effect of this allocation of machine capacity is shown in Fig. 3. The bars represent the total available time on each machine group. The shaded portions are the utilized time.

Fig. 4 shows the inventory accumulation that results from this program. Inventory is accumulated in the first and second quarters and then dissipated in the higher demand of the third and fourth quarters. The peak accumulation is the product of 1100 hr of production. The total cost of carrying the accumulation is $2100. This is the direct cost of leveling the work load.

The basic cost data used in the calculation are relative. To evaluate the program completely, it is necessary to compare it with some other program.

Assume a program in which the demands are supplied as they occur and parts are purchased to accommodate the peak demands in the third and fourth quarters. Such a program is shown in Table 3. In this program no inventory is carried from one quarter to another. The production schedule for any one quarter includes only parts that are required in that quarter. This program, however, results in 750 hr of layoff in the first quarter and 250 hr of layoff in the second quarter.

By comparing the cost of the previous program with this one, the cost of eliminating these layoffs can be determined. This comparison is given in Table 4. The costs of carrying inventory, of alternate routings, and of unprofitable purchases are shown; the savings from profitable purchases are deducted to show the net cost of each program. The difference in net costs is $1650. This is the actual cost of leveling the work load.

TABLE 4 COMPARISON OF PROGRAM COSTS

	Program without leveled work load	Program with leveled work load
Cost of carrying inventory accumulation	None	$2100
Cost of alternate routings	$1900	1900
Cost of unprofitable purchases	50	None
Total	$1950	$4000
Savings from profitable purchases	$2750	$3150
Net	$ 800 savings	$ 850 cost
Hrs of layoff	1000	None
Cost of leveling work load: $1650 to eliminate 1000 hr of layoff		

The model problem also can be used to illustrate the value of by-product information. An example is the capacity of machine group B. The calculated program provides for 100 per cent utilization of this group throughout the year. Management may wish to consider the possibility of adding group B capacity at the expense of the capacity of some other group. The value of this can be quickly determined from the work sheets of the Modi calculation. Extra group B capacity at the expense of group C capacity will provide a net gain of $1 per hr up to 300 hr per quarter. Thus making available 300 additional hours of group B time will accomplish a net saving of $1200 per year.

Many other facts can be obtained by adjusting the data and recalculating the program. Some of the decisions for which supporting facts can be developed are:

1 Whether to obtain additional capacity.

2 Whether to change the present manning plan, i.e., one man on four machines.

3 How to prepare for problems that will present themselves if the forecast is high or low.

Results of the Case Example

The solution of the model problem demonstrates the methods used for the case example. Considering a larger number of parts merely makes a longer Modi matrix. Since much of the clerical work of preparing data is done with punched cards, this is no great problem.

In the actual problem as in the model, linear programming shows that it is necessary to start overproducing some items as early as the first quarter if the seasonal peak is to be met. Inventory accumulation this early in the year was never considered under previous planning methods. Management was reluctant to allow excess capacity in slack seasons to be used for producing parts ahead of the time that they were needed. People were laid off in slack seasons despite the fact that the seasonal peak usually made overtime necessary later in the year.

Now, the factual information provided by linear programming provides a firm basis for planning a stable work load. Each calculated program shows specifically how each part is produced or procured to provide the desired work load at minimum cost.

The cost of leveling the shop work load is proving to be much less than management had anticipated—about $5000 per year. The beneficial effects of improved employee relations are already noticeable. In addition, management does not face the prospect of calling furloughed employees back to work only to discover they have found other employment and are not available.

By-products of the calculations have been found useful in planning the acquisition of new machines and in determining whether or not overtime operation will be profitable in periods of peak demand.

With punched-card equipment preparing the data, new and revised programs can be set up and calculated quickly. Forecast revisions are readily taken into account.

During the most active period at the end of the year, it is possible to program the production and procurement by months instead of quarters. This refined program can then be revised each month to assure prompt adjustment to changes in demand and complete dissipation of the inventory accumulation.

Conclusion

In this case example, linear programming has been shown to be a valuable and practical tool for management planning. The technique demonstrated in this paper will be valuable in similar situations in many other companies.

This example is typical of the way in which modern techniques of scientific management can provide factual and quantitative data on which to base decisions.

REFERENCES

1 "Industrial Engineering Handbook," by H. B. Maynard, McGraw-Hill Book Company, Inc., New York, N. Y. (forthcoming).

2 "Linear Programming in Production Scheduling at SKF Industries, Inc.," by R. O. Ferguson and L. F. Sargent, Carnegie Institute of Technology Proceedings, May, 1953.

3 "The Use of Linear Programming for Production Planning and Control," by R. O. Ferguson, paper presented at AMA Operations Research Seminar, September 1, 1954.

4 "Linear Programming and Inventory Management Seminar Proceedings," Methods Engineering Council, Pittsburgh, Pa., Sept., 1955.

A Combustion System for Spark-Fired Gas Engines Using Diesel Compression Ratios

By L. D. THOMPSON,[1] R. H. BEADLE,[2] and F. A. BLAKE,[3] BELOIT, WIS.

"Jet ignition" is a combustion system which eliminates the requirement of air modulation with varying loads in a spark-fired gas engine. Jet ignition coupled with diesel compression ratios results in a versatile engine having high thermal efficiency and simple controls. Engine governing with jet ignition only requires control of the fuel as in diesel engines. The paper describes this combustion system and the background leading to its development.

Introduction

IMMEDIATELY after World War II, the decision was made to embark on an extensive engineering program for the purpose of broadening the use and application of the opposed-piston (OP) line of engines. One of the phases of this program was the development of a dual-fuel model of the $8^1/_8 \times 10$ OP engine, which uses natural gas as a major fuel. Successful dual-fuel operation naturally led to a "spark-ignited" gas-engine development program.

The basic object of the spark-ignition program is to retain the thermal advantages inherent in high-compression gas engines while eliminating the liquid-fuel requirements. Experience has shown that while the quantity of liquid fuel for dual-fuel engines may be only 5 per cent of the full-load fuel the cost may be as high as 30 per cent of the total fuel bill.

Opposed-Piston Engine

The background of the jet-combustion system is best introduced with a description of the engine to which it was first applied. This is the OP engine, a cross-sectional view of which is shown in Fig. 1. It is a 2-cycle blower-scavenged engine having its air ports near the top of the liner and its exhaust ports near the bottom. The bore is $8^1/_8$ in. and the combined stroke is 20 in. It is a two-crankshaft engine. Each crank has a stroke of 10 in. making it an $8^1/_8 \times 10 \times 10$ engine.

In the diesel and dual-fuel models liquid fuel is injected through two nozzles mounted on the side of the cylinder approximately midway between the air and exhaust ports. In the gas models (dual-fuel and spark ignition) a gas-admission valve is located in the same horizontal plane as the injection nozzles and opens directly into the cylinder.

The Dual-Fuel Cycle. Fig. 2 reviews the major events in the dual-fuel cycle which uses natural gas as the major fuel and fuel oil for ignition:

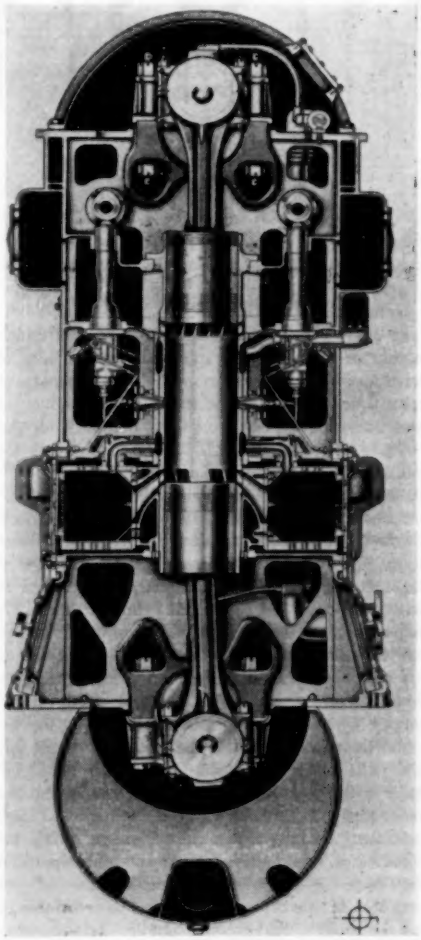


Fig. 1 Transverse Section—Model 38DS-$^1/_8$ Engine

1 Start of compression—air and exhaust ports closed—admission of gas before compression has progressed far enough to require high gas-supply pressures.

2 Compression of air and gas fuel to about 525 psi gage.

3 Ignition—just prior to inner dead center a small quantity of liquid fuel is injected which ignites the gas charge.

4 Burning and expansion of the gas-air-fuel-oil mixture.

5 Exhaust of burned charge and scavenging of the cylinder before the next cycle.

This opposed-piston 2-cycle diesel and dual-fuel design was used in developing the new spark-fired combustion system. The continuous unturbocharged industrial rating of the engine is 85 bmep at 720 rpm or 160 bhp per cylinder.

---

[1] Superintendent, Experimental Department, Fairbanks, Morse & Company.

[2] Assistant Superintendent, Experimental Department, Fairbanks, Morse & Company.

[3] Project Engineer, Experimental Department, Fairbanks, Morse & Company.

Contributed by the Oil and Gas Power Division and presented at the Oil and Gas Power Division Conference, New Orleans, La., April 1–5, 1956, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, February 6, 1956. Paper No. 56—OGP-6.

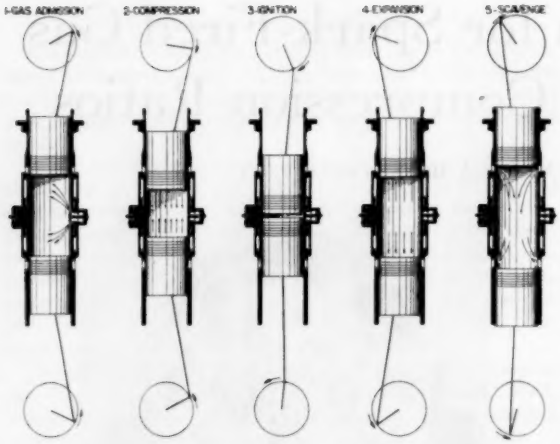


FIG. 2 MAJOR EVENTS IN THE DUAL-FUEL CYCLE

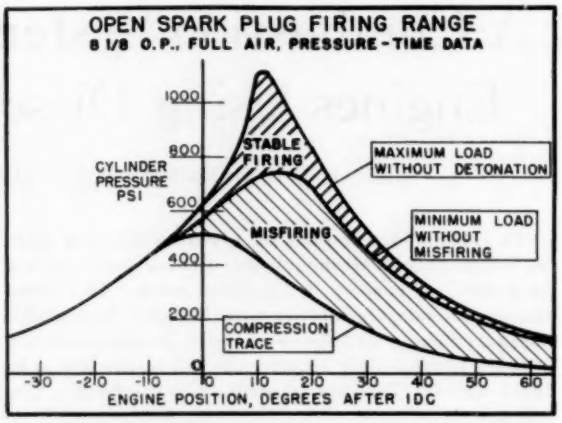


FIG. 3 OPEN SPARK-PLUG FIRING RANGE—PRESSURE-TIME DATA

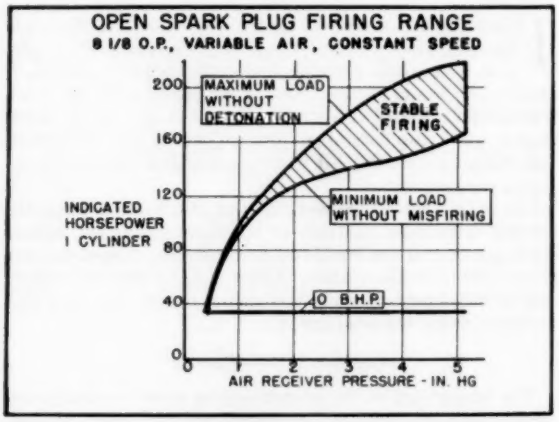


FIG. 4 OPEN SPARK-PLUG FIRING RANGE—VARIABLE AIR

TESTING SPARK-IGNITION SYSTEM

The first engine modification for testing spark ignition consisted merely of replacing the injection nozzles of a dual-fuel engine with spark plugs and providing the necessary electrical system for a timed spark. As the experimental engine was a relatively large unit, 960 hp, requiring large quantities of gas, it was decided to proceed cautiously. Thus the exploratory tests were conducted on only one of the engine's 6 cylinders.

The test cylinder was equipped with a separate gas supply which could be varied independently from the other five cylinders. With this divorced cylinder arrangement, the engine was started and brought up to rated speed with the five diesel cylinders. When the engine steadied out with reference to speed and temperature, gas was admitted to the test cylinder. The main object of the single-cylinder test exploration was to determine the limits within which the spark-ignited combustion was regular. Regular firing was checked by a point-by-point pressure-time indicator. A sample of these results is shown in Fig. 3.

The engine operating conditions for the first test were the normal conditions for the dual-fuel engine and the air supply was not varied with load. The lowest load which could be carried with regular firing was determined by increasing the gas to the test cylinder until the expansion side of the card was free of a no-fuel-cycle trace. As long as the engine was misfiring, the point-by-point indicator would show several different expansion traces. A complete misfire would trace the same as a no-fuel cycle.

The next step of this type of testing was to determine the maximum load the spark cylinder could handle. In this case the procedure was to increase the gas supply until either detonation or autoignition occurred, back off slightly, and take another indicator card.

The indicated-horsepower determinations made from the indicator cards showed the full air minimum and maximum load points to be roughly 75 and 100 per cent, respectively, of the commercial rating. Regular firing was only obtained in this narrow high-load band. As the gas pressure was reduced toward light load the cycle became very erratic to the point of cutting out altogether.

That the plugs fired the gas regularly near maximum load indicated that satisfactory ignition was being obtained with the richer mixtures. The next step in the single-cylinder program was then to vary the air supply. The objective was the achievement of richer and more ignitible mixtures at lighter loads, the enrichment being obtained by reducing the air quantity. This was done by installing a bypass valve on the scavenging blower and adjusting it to give various reduced air-box pressures. With this variable air system, another series of maximum and minimum load determinations was made for the various air settings. The results of these tests are shown in Fig. 4. With the blower bypass it was possible to vary the air-box pressure over the range of 0.4 to 5.2 in. Hg. It is realized that this index does not give an exact indication of the actual pounds of air trapped in the cylinder. However it is a practical way of changing the air quantity and conducting reproducible tests.

The curves show maximum and minimum indicated horsepower that could be obtained regularly at the various air-box pressures. The shaded area enveloped by the maximum and minimum curves indicates the horsepower range over which regular firing could be expected for various air quantities. To put it another way, this area shows the range within which the air to the cylinder would have to be varied with load to obtain regular and safe operation. The stable firing range became progressively narrower as the air and loads were reduced. In fact, the range might be described more accurately as a point for the light loads.

The existence of this narrow stable-operation range was quite accurately verified at a later date when all six cylinders of the test were converted to the open-chamber spark system. The test

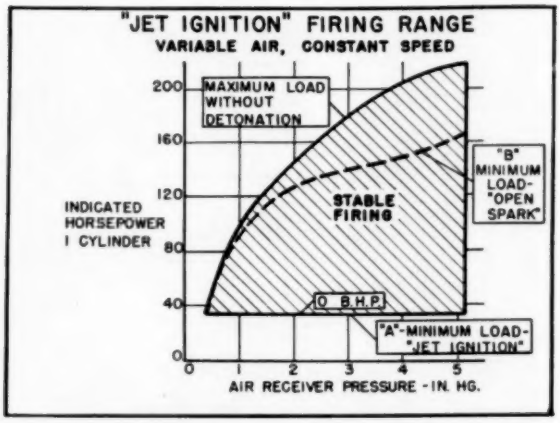


FIG. 5 JET-IGNITION FIRING RANGE

showed that once the engine was nursed through the light-load range to approximately 75 per cent load, stable operation was obtained. Getting the engine through the light-load range however was no simple task.

RESULTS OF EXPLORATORY TESTS

The results of the exploratory tests indicated that two avenues of development were open for obtaining a safe and versatile engine. One would be to develop a governor-actuated control system which would keep the air-fuel mixture in the proper range. The other approach would be to widen the operating range.

The second avenue, widening of the combustion range, was by far the most attractive as every step of progress in its development would reduce the complexity of controls. Complete success of the widening approach would result in a combustion range as shown in Fig. 5. Here the open-plug minimum-load curve B has been dropped to curve A which represents zero load. With such a widened combustion range there would be no need for automatic air-fuel-mixture controls. The widened range would be more effective than even the most complex controls in eliminating misfiring, an item which it is most desirable to eliminate because of the accumulation of raw gas in the exhaust system which usually follows. This raw-gas accumulation is very serious with large engines owing to the system size and gas quantities.

The complexity of controls which would control within the open-plug stable range was enough to push this approach into the background. Variables which would have to be handled would include speed, load, barometric pressure, and ambient temperature. It has been our experience that controls of this nature, requiring correction for so many variables, frequently are misunderstood and then improperly adjusted. This leads to unsatisfactory performance of the product and even at times unsafe operation.

It was due to the basic desire to achieve the maximum simplicity and safety that the decision was made to pursue development along the second avenue, to work on combustion itself to widen the operating range and thus eliminate the need for fancy controls.

Table 1 gives a picture of air-fuel ratios involved in these diesel compression engines. Very lean mixtures—32:1 at maximum load down to 63:1 at quarter load—are regularly ignited in dual-fuel engines by liquid pilot oil. With pilot-oil ignition modulation of the air to eliminate misfiring at low loads is not required. This small quantity of liquid fuel apparently liberates sufficient energy to ignite these very lean mixtures. The table also includes a column showing fuel-air ratios which would be expected in a spark-fired gas engine performing with the same thermal efficiency as a dual-fuel unit. Spark engines show richer mixtures than dual fuel because of the extra gas needed to replace the liquid fuel. The richest mixture, occuring at rated load, would be 29:1. Mixture strength would reduce to something less than 53:1 for no load or idle operation.

TABLE 1 APPROXIMATE TRAPPED AIR-FUEL RATIOS (LB AIR/LB FUEL)

	Dual-fuel engine	Spark engine
Engine load % rating	Air-gas mixture ignited by pilot oil	Air-gas mixture, no pilot oil
100	32:1	29:1
75	42:1	38:1
50	54:1	47:1
25	63:1	53:1

A number of ideas were tested in an attempt to fire these extremely lean mixtures. Included were such items as one or two plugs per cylinder; various spark-plug designs; different piston-top contours, and lower compression pressures. The low-compression test in which the pressures were reduced from 525 psi to 350 psi showed no improvement whatsoever in broadening the ignitibility range. In fact, none of the foregoing items showed any promise in obtaining a more versatile spark-fired engine.

DEVELOPMENT OF JET-IGNITION SYSTEM

At this stage it was concluded that the triggering energy from a spark plug, alone, would not be sufficient to fire these very lean partial-load charges regularly. It was known, however, that these lean charges can be regularly and safely fired in dual-fuel engines where pilot oil releases relatively large quantities of energy for ignition purposes. From this line of reasoning, it was considered that the best approach would be to use the spark for igniting only a small protected and enriched localized charge. The energy released from ignition of this localized charge would be many times greater than from the spark alone and probably would ignite lean mixtures as pilot oil does in dual-fuel engines. Protection of the charge is mentioned as well as enrichment. Some evidence from early tests indicated that the spark was firing the lean charges, but the resulting flame was being extinguished or quenched before consuming much of the main charge.

A new combustion-system design, based on this theory of providing a protected and enriched localized charge around the spark plug is shown in Fig. 6. This shows the basic design arrangement of the FM jet-ignition system.

The main gas charge, throttle-controlled by the governor, is admitted to the combustion chamber through a poppet valve.

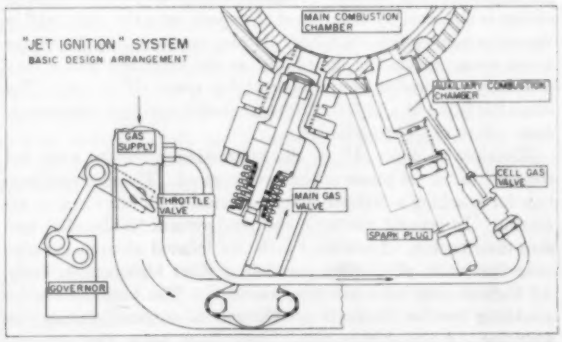


FIG. 6 BASIC DESIGN ARRANGEMENT OF FAIRBANKS, MORSE & COMPANY—JET-IGNITION SYSTEM

This is the same basic gas-control system as used in the dual-fuel models. The gas-supply poppet valve is opened by a cam for a short period early in the compression stroke. The spark plug is located in an auxiliary combustion chamber. The auxiliary chamber communicates with the main chamber through an orifice. The auxiliary chamber has an additional gas-supply source which is not varied by the governor. This is called "cell" gas and its supply line is maintained at a constant pressure regardless of engine speed or load. The extra gas is admitted to the auxiliary chamber through an automatic one-way check valve. When the cylinder pressures are low, during the exhaust and scavenging periods, gas flows through the nonreturn valve to the small chamber containing the spark plug.

In this manner the enriched localized charge was obtained. Referring back to the early tests it was shown that regular spark firing could be expected with mixtures as lean as 38:1. Now with this extra cell-gas arrangement it is easy to see how the mixture surrounding the spark plug can be enriched to as much as 14 to 1; 14 to 1 is the trapped-volume compression ratio of the engine, and such a mixture would result from a cell full of gas at the start of compression if there were only air in the main chamber.

An air-gas mixture of 14:1 is certainly too rich a charge for these high-compression engines to handle without autoignition prior to the spark. It does show, however, that by proper sizing of the flow passages in the cell supply line it is quite feasible to have mixtures surrounding the spark plug richer than 38:1 for all throttle requirements, zero through full load.

How System Functions. The jet-ignition system functions during each engine cycle as follows: (1) It automatically obtains a readily ignitible charge in the auxiliary chamber; (2) the spark fires the localized auxiliary charge; and (3) the combustion which follows in the auxiliary chamber blasts a jet of flame through the orifice into the main chamber where the jet ignites the main charge.

The very first tests with this system showed a definite improvement in light-load operation. Subsequent experiments with various proportions of the auxiliary-system components resulted in consistent and regular ignition throughout the entire speed and load range when using diesel compression ratios. This versatile operation is obtained without modulation of the air supply. As with the dual-fuel engine modulating controls for varying the air supply to improve part-load efficiency are optional. These controls, however, are not required to eliminate misfiring or erratic operation.

Results of Jet-Ignition Tests. The refinement experiments on the jet-ignition components determined the effects of variable cell volumes, orifice areas, and quantities of cell gas. The most startling outcome of this testing was the elimination of the cell-gas requirement in one size of engine. It appears that the ignitible charge is obtained by isolation of the spark from the high swirl in the main combustion chamber. It also appears that the engine speed range affects the ignitibility as the extra cell gas is very definitely required in the smaller higher speed OP engine. The small OP is the $5^1/_4 \times 7^1/_4$, 1200-rpm model having a continuous-duty rating of 75 hp/cylinder.

Electrical System. All of the interesting problems were not confined to the jet phase of the development. The electrical system for sparking a 500-psi charge demanded its fair share of attention. In general the basic electrical system features low-tension distribution. Individual coils are located at each cylinder near the spark plugs thus permitting short high-tension leads. All high-tension leads are fully shielded. The basic reason for shielding was to eliminate all chances of external shorting or sparking. A suitable primary low-voltage pulse was obtained with various types of conventional equipment. Low-tension magnetos, battery breaker-point systems, and pulse generators

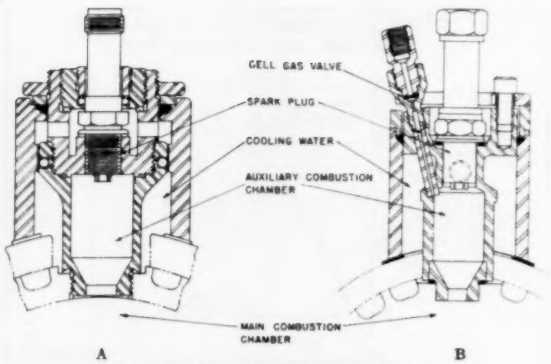


FIG. 7 JET-IGNITION COMBUSTION DESIGNS
A, with spark-plug water cooling
B, without spark-plug water cooling

were tried and all provided sufficient spark. Initially the pulse generator had the edge on the other systems with respect to simplicity and durability. Consequently all production engines at the present time are equipped with these pulse generators. During the past months there have been interesting new developments in the low-tension magneto field which in the future certainly will provide good healthy competition for the pulse generator.

Spark Plugs. The spark plug itself presented some interesting problems. As mentioned before, the very first test with the jet system gave more positive ignition. The cell was working so well, in fact, that after 3 min of heavy-load operation the entire spark plug turned red. The problem of keeping the spark-plug temperatures within a usable range was solved by leaning the mixture in the cell and modifying the cell designs for better cooling of the plug. One design which is performing very well involves direct water cooling on the plug as shown in Fig. 7(A). More recent developments indicate that possibly the direct cooling can be eliminated with new cell designs similar to Fig. 7(B). As to the various types of plugs tested, the Champion R115 gave the best performance and is being used in both the $5^1/_4$ and $8^1/_8$ production engines.

MOVING INTO THE FIELD

At this stage, the basic objective of being able to spark fire a high-compression gas engine throughout its entire load range without the need for modulating air controls had been satisfied in the laboratory engine. The next step was to move into the field for studying the customer acceptance and durability aspects. Arrangements were made with the municipal officials of Stockton, Kan., to apply the jet-ignition system to the 6-cylinder $8^1/_8 \times 10$ OP engine in their municipal light plant. Alternating-current power generation for municipal light plants is a very exacting service. Engines in this service are required to operate over wide load ranges and yet maintain extremely accurate speed regulation. At night, even though the electrical loads are very light, speed control must be accurate to keep time for all the electric clocks.

Fig. 8 shows the Stockton installation. The Stockton unit was first converted to spark operation in September, 1952, and has now logged more than 19,000 hr in this exacting municipal service. During the 19,000-hr period, the maintenance interval for spark plugs was 1800 to 2400 hr. The plugs could be regapped twice, therefore resulting in a plug life of more than 5000 hr. To the best of our knowledge this is the first successful spark-fired 2-cycle gas engine in a-c generating-service operation. The American Bosch pulse generator and high-voltage transformers were trouble-free throughout this entire period.

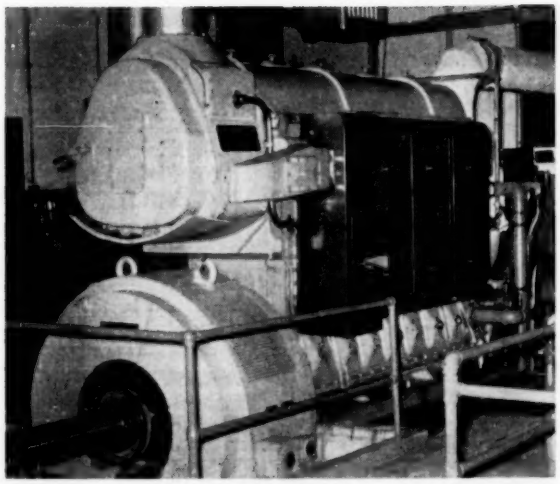


FIG. 8 STOCKTON, KANSAS, INSTALLATION

are repeatedly operated from idle to full torque, full speed, and with either propane or natural gas, certainly shows that the jet system will fire regularly over a wide range of operating conditions.

FUTURE OF JET-IGNITION SYSTEM

Concluding this presentation we think it is appropriate to take a look into the crystal ball for hints as to the possible future of the jet-ignition system. First, the already established good performance of jet ignition in a-c generating plants opens a new field for reducing fuel cost. In areas where gas is readily available it no longer will be necessary to use both liquid and gas fuels to retain the high efficiencies afforded by diesel compression ratios. In fact, even in areas where the gas supply may be interrupted temporarily, propane could be used as a standby in these emergencies and thereby eliminate the continual need for relatively expensive liquid fuel. Many installations now pay thousands of dollars each year for diesel fuel required by dual-fuel engines just to protect against a gas interruption. Experience has shown that these interruptions usually average not more than two days each winter. Jet ignition with simple propane standby facilities should handle this situation nicely.

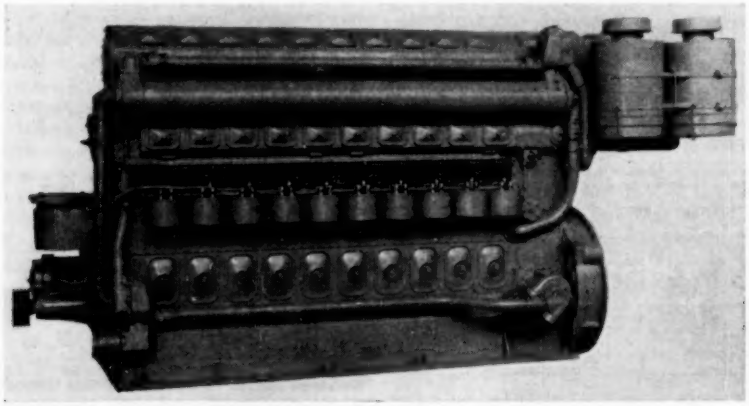


FIG. 9 5¼ OP SPARK-IGNITION ENGINE

Economic Considerations. As to economics, the jet-ignition system in this particular installation showed a 19 per cent reduction in fuel cost as compared to the same engine operating on the dual-fuel cycle using liquid pilot oil for ignition.

ADDITIONAL SPARK MODELS INSTALLED

To illustrate further the combustion regularity and the over-all reliability of the jet-ignition system, several additional 8⅛ × 10 OP spark models have been installed in other municipal generating plants. All of these units are maintaining the necessary accurate speed control over the complete load range. The first of these new production installations logged over 7500 hr during its first 11 months of operation, and 7500 hr in 11 months means pretty continuous service.

As previously mentioned the jet-ignition principle also was applied to the FM 5¼ × 7¼ OP gas engine. Fig. 9 shows the small 5¼ equipped with a pulse generator and coils for jet ignition. The first spark-fired small OP's were all applied to oil-field drilling rings. The No. 1 rig installation has now been in service approximately 7500 hr and has completed the 16th hole ranging between 7000 and 8000 ft in depth. The rig operated on either propane or natural gas. Oil-field-drilling service, in which engines

Finally, jet ignition's already established fine performance with propane fuel makes it feasible to consider spark-fired high-compression gas engines for mobile equipment such as switching locomotives, towboats, and so on. The thought of propane locomotives becomes even more intriguing when one considers the current emphasis on smoke abatement.

A new approach to the old problem of operating spark-ignition engines on a high-compression cycle was made with jet ignition. The result was not only an engine with diesel compression but one which did not need any of the traditional carburetion controls so long associated with spark engines. That engines using the jet system measure up to being reliable, stable, and versatile is evident in the field records.

Discussion

E. G. BEARDSLEY.[4] This discusser believes that the engineers of every company who have built, or are now building, large two-cycle gas engines have at one time or another made an effort to improve the regularity of firing these engines at partial loads.

[4] Vice-President, Engineering, Clark Brothers Company, Olean, N. Y. Mem. ASME.

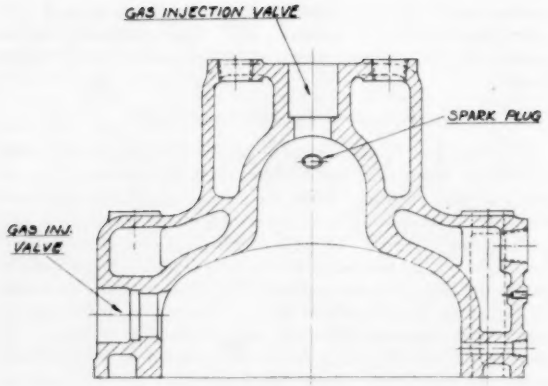


FIG. 10

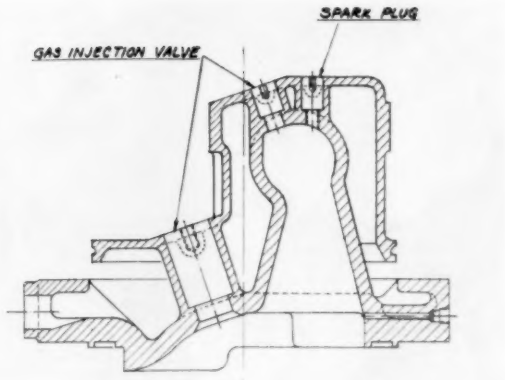


FIG. 11

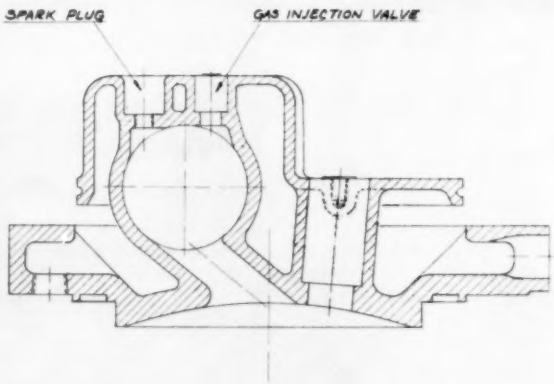


FIG. 12

It is quite probable that many of them have come up with a similar idea to that which is presented in this paper, although the actual execution in the form of a mechanical design may have been somewhat different.

Two engineers working for this discusser's company did have this same idea working entirely independently a short time after they attacked the problem. In other words, if the engines misfire at partial loads due to the mixture being too lean why not divide off part of the combustion chamber so that when enough gas is injected into the small chamber to operate the engine at no load, only enough air will be present to give a good firing mixture?

Fig. 10 shows the first design that we built and tested. This was installed on a horizontal 16×20 single-cylinder gas-engine-driven compressor unit. The light load firing was remarkably good but the head unfortunately developed a crack in it and the project was abandoned as there was not sufficient interest to carry on further work at that time.

Fig. 11 shows another approach to this problem, this head having been installed in 1940 on a model RA 14×14 angle-type gas-engine-driven compressor unit. Once again greatly improved firing resulted but the maximum load operation was poor.

Fig. 12 shows another design which was installed on an engine in 1941. This gave extremely good results on a single-cylinder unit, but was not as good on a multicylinder machine presumably because of uneven air distribution and scavenging which always occurs on the multicylinder engine to some extent.

The three combustion chambers described in the foregoing were used with low compression cylinders. The compression pressure was about 110 psi in each case. The percents of total combustion-chamber volume in each case were 12 per cent, 6.5, and 16 per cent, respectively.

One will notice that on all three designs a small gas-injection valve and a spark plug are provided in the auxiliary chamber as well as the main injection valve in the main part of the chamber. This project was abandoned because of the press of war work and although work on other methods of improving light load firing have been investigated and some adopted, this project was never taken up again.

The authors are to be complimented on their paper and on the work which they have done on this project.

W. R. CROOKS.[5] Messrs. Thompson, Beadle, and Blake are to be complimented for obtaining, and reporting on, a simple solution for a difficult problem. The Cooper-Bessemer Corporation has done extensive work along this line, with comparable results. The conventional two-cycle gas engine, whether it be low compression—100 psig, medium compression—250 psig, or diesel compression—from 400 psig up, is known to tend to miss at light loads due to the extremely lean air-fuel mixture. The jet ignition presumably supplies the extra energy required to ignite the lean mixtures at part loads.

It seems likely that engines operating with this type of ignition will have the same type of difficulty we encountered in the early days of gas-diesel operation. For a short period of time we operated the gas-diesel, or dual-fuel engines with the same amount of air we use for diesel operation. As the load decreased, the air-fuel ratio became increasingly leaner until the limits of inflammability were reached which agree with the results the authors have reported.

With the pilot fuel set at 10 per cent of the total full-load fuel, we obtained the following one-quarter load performance: (*a*) High pilot fuel-oil requirements to ignite the very lean mixtures. (*b*) High gas fuel consumption. (*c*) Objectionable exhaust odors. The exhaust gases were obnoxious from the odors of aldehydes and were sufficiently strong to irritate the mucous membranes of the noses of persons in the vicinity of the engine operation.

It appears as if the combustion was so slow in some portions of the cylinder that combustion never proceeded beyond the "cold-flame" stage of combustion before it was discharged into the exhaust. It would seem as if these conditions would prevail with jet-ignited partial-load operation of two-cycle engines.

[5] Cooper-Bessemer Corporation, Mt. Vernon, Ohio.

A. W. Hussmann.[6] It may be of interest to note that the very same idea of jet ignition has been in the air for quite some time in connection with the gasoline Otto-engine. We have the same problem of misfiring for lean mixtures in our cars. There we provide for the proper air-fuel ratio by throttling the air for part loads. This throttling is very wasteful and is, to a large extent, responsible for the shamefully poor fuel economy. We are putting more and more horsepower in our cars with the result that we are always working with throttling in the uneconomical part-load range.

The jet-ignition system seems to be one answer to the problem of how to operate our spark-ignition engines without air throttling, i.e., with good part-load economy. With fuel admission by carburetion, this was not practicable. However, with the advent of gasoline injection, the idea seems to be highly promising. And it is more than just an idea. Promising test results have been published by Prof. J. J. Broeze[7] of Delft, Holland, and Prof. W. E. Meyer of The Pennsylvania State University has done some preliminary testing with an unthrottled gasoline-injection engine with most encouraging results. Here is definitely a field where further research and development are needed.

W. K. Newcomb.[8] This interesting paper tells how a difficult problem was solved. The auxiliary combustion chamber or cell where an easily ignitible gas and air mixture is maintained is a very ingenious device. If it permits smooth operation at part loads, it has greatly improved the flexibility of a type of engine where modulation of the air is difficult and not very successful.

Four-cycle engines, on the other hand, are inherently easy to control and their smoothness of operation and flexibility are well known. With naturally aspirated four-cycle spark-ignition engines the constant-quality mixing valve is simple and reliable. With turbocharged four-cycle engines the air control is also simple. Here the speed of the exhaust-gas-driven supercharger responds to the load and this raises or lowers the manifold air pressure. However, we find this control is not enough and also use a butterfly valve which is positioned by the air or gas pressure. With a centrifugal blower we can do this without having to blow off any air because the blower operates farther back on its characteristic curve.

The authors imply that air modulation improves the fuel consumption at part loads. It would be interesting to know what part-load fuel consumption they get without air modulation.

Table 1 mentions 53:1 air-fuel ratio. This sounds more like total air-to-fuel ratio without allowance for scavenging. How were this and other mixture ratios measured?

Authors' Closure

We certainly appreciate the fine discussions and wish to thank the discussers for their interest and efforts spent in the preparation of their comments.

Mr. Beardsley's experiences with low-compression gas engines utilizing divided combustion chambers is quite interesting. His basic experiments parallel our early work in being directed toward enrichment of a small section of the trapped combustion charge. This is added evidence that the principles involved in the jet-ignition combustion system are not limited to any one engine. The engine details between the two programs vary widely but their reactions toward improved light-load operation with the local enrichment are similar.

Mr. Crooks implies that he would expect the operation at light loads to be erratic regardless of the ignition system so long as the air supply is not reduced below that required for the same engine operating as a diesel oil engine. We definitely disagree with this in the case of our dual-fuel and spark-ignition engines which both normally use the same air as their diesel counterpart. We have never reached a pilot-fuel limit below which combustion became irregular in the dual-fuel engine, but rather are limited on pilot-fuel quantity by the least amount we can meter consistently. Jet ignition gave consistent and regular ignition throughout the entire speed and load range. Air modulation or reduction controls are optional where reducing the light-load fuel consumption is important.

Objectionable exhaust smells have never been reported around the engines although none has been applied in tightly enclosed applications such as mines.

Professor Hussmann's comments regarding the application of jet-ignition principles in gasoline engines further points out the wideness of the field of potential applications for the jet-combustion system.

Mr. Newcomb requests clarification of the method for determining the stated air-fuel ratios and for the values of part-load fuel consumptions on spark-ignition gas operation when the air supply is not reduced.

The air-fuel ratios were determined on a per cycle basis using the pounds of air trapped in the cylinder at air-port closing divided by the pounds of fuel used per cycle. The 50 per cent rated-load fuel consumption without air modulation for the larger opposed-piston engine is approximately 12 per cent above that obtained with reduced air.

Dr. Elliot raised questions on lean charges, adjustments for propane and natural gas, odor in exhaust gases, and light-load operation of dual-fuel and spark-fired gas engines, while making an eloquent plea for more complete data from development programs.

By "firing of lean charges" we mean obtaining a regular combustion cycle with the lean charges. As he indicates, unburned gases are present after combustion of these charges. The burning of these unburned gases is the major source of reduced fuel consumption with reduced air, which is optional, when operation at these loads is prolonged.

The fuel consumption on a Btu basis is very close on propane and natural gas. Owing to the differences in the specific heating values, this results in lower gas-system operating pressures at all points when using propane. The energy released by the cell is very close for both fuels.

Exhaust odor is treated in the comments on Mr. Crook's discussion. Light-load combustion with full air is stable with both dual fuel and jet ignition. It was the knowledge that the pilot oil regularly ignited these extremely lean mixtures that spurred on the jet-ignition development.

---

[6] Professor of Engineering Research, The Pennsylvania State University, University Park, Pa.

[7] "Automobiles and Petroleum: Past, Present and Future," by J. J. Broeze, The Institution of Mechanical Engineers, London, England, 1953/54, No. 7.

[8] Mechanical Engineer, Ingersoll-Rand Company, Painted Post, N. Y. Mem ASME.

The Application of an Analog Computer to the Measurement of Process Dynamics

By P. E. A. COWLEY,[1] EMERYVILLE, CALIF.

The usual method of obtaining frequency-response measurements is described. The data obtained are reduced by visual inspection, but difficulties arise due to distortion and noise in the process. The range of frequencies over which measurements can be made is limited by noise and distortion and may prove inadequate for studying the stability of the process when controlled. These shortcomings may be overcome by methods of frequency-response measurement in which the data are reduced by means of a computer rather than by visual inspection. Two methods of frequency-response measurement using an analog computer for data reduction are described. Results are given of laboratory measurements on a dummy process and of field measurements on a real process. The advantages of computer reduction of the data are clearly demonstrated. The further advantages to be derived from digital computation rather than analog computation are indicated.

Nomenclature

The following nomenclature is used in the paper:

A = amplitude ratio of process-transfer function
B = amplitude of process response
C = controlled variable
D = amplitude of sine-wave disturbance
E_b = limit voltage
f_1 = a particular function of time or a sine wave
f_2 = a particular function of time or a process-time response
f_3 = a cosine wave
f_4 = a sinusoidal wave of adjustable phase
f_i = input signal
f_o = output signal
k = negative damping factor and potentiometer ratio
m_b = correction coefficient
N = integral number of periods
n = an integer
P = control signal to process
p_a = potentiometer ratio for phase null test
p_b = potentiometer ratio for amplitude measurement
R = controlling variable or set point
T = integration time or period of sinusoidal wave
T_0 = time constant
T_1 = time constant
T_2 = time constant
t = time
x = independent variable
α = phase angle between f_3 and f_4
θ = phase angle of process-tranfer function
τ = independent variable of the correlation functions
τ_1 = a delay time
ϕ_{11} = autocorrelation function of f_1
ϕ_{12} = cross-correlation function between f_1 and f_2
ϕ_{21} = cross-correlation function between f_2 and f_1
$\bar{\phi}_{12}$ = modified cross correlation between f_1 and f_2
$\bar{\phi}_{22}$ = modified cross correlation between f_3 and f_2
$\bar{\phi}_{42}$ = modified cross correlation between f_4 and f_2
ω = angular frequency of sinusoidal wave

[1] Engineer, Shell Development Company.

Contributed by the Instruments and Regulators Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS and presented at the joint ASME-ISA Conference, New York, N. Y., September 17–21, 1956.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, June 19, 1956. Paper No. 56—IRD-20.

Visual Method of Frequency Response Measurement

Usually frequency-response measurements are made as follows:

(*a*) A sine-wave disturbance is applied to the process.

(*b*) Both the disturbance and the process response are recorded on the same chart.

(*c*) The process-response amplitude and phase shift relative to the disturbance are determined from the chart record by visual inspection. This method will be referred to as the "visual method."

A typical chart recording is shown in Fig. 1. Shortcomings of this method are as follows:

1 Distortion often occurs in the process itself. The system response then consists of a fundamental component with a number of harmonics although only the fundamental component is of importance in studying the stability of the closed loop. The fundamental component is the quantity which must be compared with the input disturbance, but because of the distortion it is difficult to obtain its amplitude and phase relationship by visual inspection of the chart records.

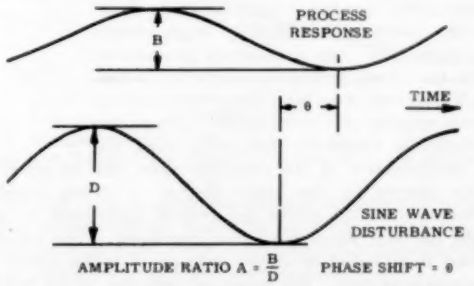


FIG. 1 TYPICAL CHART RECORD ILLUSTRATING VISUAL METHOD OF FREQUENCY-RESPONSE MEASUREMENT

2 In the process itself, there is often considerable noise which may obscure the process response. Fig. 2 shows a recording in which the process response is completely obscured by noise. The noise is a serious limitation on the frequency-response measurements especially at the higher frequencies where the response has a very much smaller amplitude than at the lower frequencies. If the phase shift of the process exceeds 180 deg, then it is desirable to extend the upper frequency of measurement at least as far as the highest frequency[2] at which the phase shift is equal to 180 deg.

[2] The phase shift may be equal to 180 deg at two or more separate frequencies. This occurs in conditionally stable systems such as temperature control of exothermic reactions.

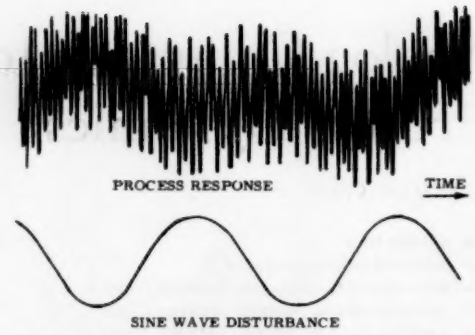


FIG. 2 CHART RECORD OF A PROCESS RESPONSE COMPLETELY OBSCURED BY NOISE

If the phase shift of the process does not exceed 180 deg, it usually becomes asymptotic to 180 or 90 deg.[3] In such cases it is desirable to extend the upper frequency of measurement as far as is necessary to determine the asymptotic value. The reason for extending the frequency-response measurements to include phase shifts of 180 deg is that, in studying the stability of the closed loop, the amplitude response at this phase angle is of great importance. In an effort to overcome the noise, it is sometimes permissible to increase the amplitude of the input disturbance, but often this is not possible in a field test because of upsetting the process itself or driving some component into saturation. The upper frequency limit at which frequency-response measurements can be made by the visual method is approximately that frequency at which the amplitude of the process response is equal to the amplitude of the noise.

FREQUENCY-RESPONSE MEASUREMENT BY CORRELATION TECHNIQUE[4]

Two methods of frequency-response measurement employing a correlation technique have been developed to overcome the limitations and shortcomings of the visual method. They are known as the "quadrature components method" and the "phase null method." Measurements may be made by these methods on processes for which the visual method is completely useless. Fig. 2 shows the sine-wave disturbance and the process response of such a process. The process response is so completely obscured by noise that the determination of the amplitude ratio and the phase shift of the process by the visual method is clearly impossible. However, these quantities are readily determined by either of the methods which have been developed. The advantage accrues from reduction of the data by computer rather than by visual inspection. For computer reduction of the test data, the data must be fed into the computer either (*a*) by locating the computer at the test site; (*b*) by recording[5] the data at the test site in such a form that it may be "played back" to the computer at the computer site. Recording may be by frequency modulation on magnetic tape for analog computation, or by teletype (or other) code punched paper or magnetic tape for digital computation; (*c*) by transmitting the data from the test site to the computer site by land line using teletype (or other) facility.

The quadrature components method and the phase null method employ the simplest concept of correlation. Provided that it is possible to apply a pure sine-wave disturbance to the "process in measurement," the computations required to apply correlation techniques to the problem may be made by a relatively simple analog computer setup which does not require a dead-time unit.

The methods as described utilize an analog-type computer. Digital computation also may be used, and will in general give superior results. It will be assumed that the computer is located at the test site, thus avoiding the question of recording or transmitting the data. The methods will first be described and illustrated with reference to measurements on a simulated two time-constant noisy process. The value of the method for the measurement of real processes will then be illustrated with results obtained on an exothermic temperature process.

EXPERIMENTAL RESULTS ON A SIMULATED PROCESS

In order to obtain a comparison of the results of frequency-response measurements by the visual method with results obtained by correlation techniques, a simulated process having two time constants was set up on the analog computer. Fig. 3 shows the computer setup. The variable-gain amplifier shown in the figure is used to amplify the process output by a known factor. When the process output is too small to be measured directly, it may be determined from a measurement of the amplified process output.

Process noise was simulated by adding a noise signal to the process output. Since a low-frequency, random-noise generator was not available, the noise signal was synthesized from four sine waves of approximately equal amplitude. Fig. 4 shows the component and the composite wave forms. The characteristics of the noise signal differed from those of random noise in two important respects; e.g., (*a*) occasional peaks of very large amplitude did not exist, (*b*) the distribution of the noise energy in the frequency range of interest was not ideal. Nevertheless, the noise signal was considered to be sufficiently representative of process noise to yield useful results.

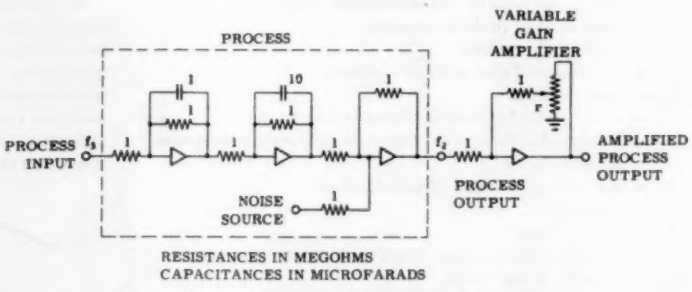


FIG. 3 ANALOG COMPUTER REPRESENTATION OF PROCESS

Frequency-response measurements were made by three methods:

(*a*) Visual method.
(*b*) Quadrature-components method.
(*c*) Phase-null method.

FREQUENCY-RESPONSE MEASUREMENTS

Visual Method. The frequency response of the process analog, Fig. 3, was measured by recording the applied sine-wave disturb-

[3] The asymptotic value is always a multiple of 90 deg for the so-called "lumped" systems. In the case of certain processes the asymptotic value may be other than multiples of 90 deg.

[4] Technical Report 117-55, July 29, 1955, Shell Development Company, not generally available.

[5] If the recording system enables playback to be made at a speed faster than the recording speed, the data reduction may be performed in accelerated time rather than in real time. Owing to the cost of digital-computer time this facility is a must for digital-data reduction.

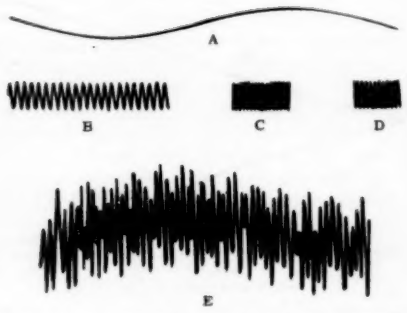


FIG. 4 NOISE COMPONENTS (*A*) TO (*D*) AND COMPOSITE NOISE WAVEFORM (*E*)

ance and the process response; i.e., by the technique currently employed in regular frequency-response measurements. The amplitude ratio and phase shift of the "process" as obtained by visual reduction of the chart records are given in Fig. 5. The theoretical curves computed from the known time constants of the analog process also are plotted in the same figure. It will be seen from the figure that the accuracy of the measurements deteriorates progressively as the amplitude ratio falls below the noise level. The accuracy is good for measurements at frequencies up to that frequency (4 cycles per min) at which the amplitude of the process response is equal to the amplitude of the noise. From the curve through the measured points it is possible to estimate the longer time constant of the process, but the shorter time constant cannot be estimated even approximately. Because the additional phase shift beyond 90 deg might be due to dead time, it is impossible to decide whether the process has two (or more) time constants, or one time constant and dead time. The highest frequency at which measurements were possible was not high enough to determine the essential dynamics of the process.

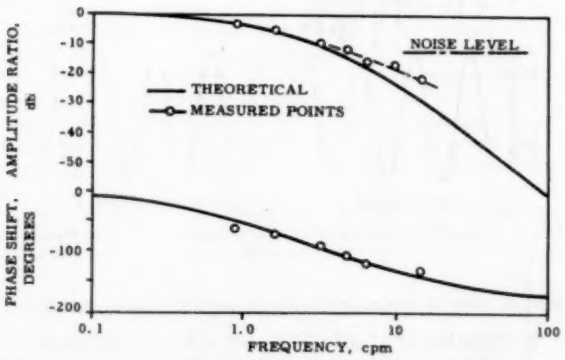


FIG. 5 AMPLITUDE RATIO AND PHASE SHIFT BY VISUAL METHOD OF FREQUENCY-RESPONSE MEASUREMENT

Quadrature-Components Method.[6] In this method the process output is analyzed as two components, one "in-phase" with the process input and the other "in-quadrature" with the process input.

The method is shown in block-diagram form in Fig. 6. The sine-wave generator which also generates cosine waves is described in the Appendix. The correlation computer computes the

[6] A similar method has been described in "Ein Verbessertes Verfahren zur Frequenzgang—Analyse Industrieller Regelstrecken," by Schäfer and Feissel, *Regelungstechnik*, no. 9, vol. 3, 1955.

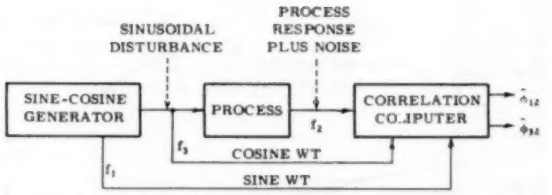


FIG. 6 QUADRATURE COMPONENTS METHOD OF FREQUENCY-RESPONSE MEASUREMENT

modified cross correlations $\bar{\phi}_{22}$ and $\bar{\phi}_{12}$ which are measures of the in-phase component and the quadrature component of the process output. The correlation computer and the derivation of the modified cross correlations are described in the Appendix.

The transfer function of the process has an amplitude A and phase angle θ which may be determined from $\bar{\phi}_{22}$ and $\bar{\phi}_{12}$

$$A = 2\sqrt{[(\bar{\phi}_{22})^2 + (\bar{\phi}_{12})^2]} \quad [1]$$

$$\theta = -\tan^{-1}\left(\frac{\bar{\phi}_{12}}{\bar{\phi}_{22}}\right) \quad [2]$$

A plot of the frequency response measured by this method is given in Fig. 7 together with the theoretical curves. Accurate measurements were possible up to a frequency of 95 cycles per min (cpm); i.e., a frequency limit 20 times greater than the frequency limit obtained with the visual method of frequency-response measurement.

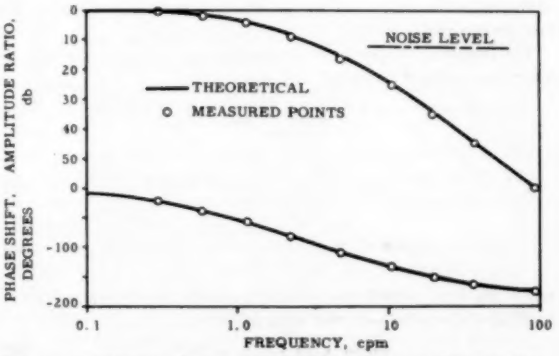


FIG. 7 AMPLITUDE RATIO AND PHASE SHIFT BY QUADRATURE-COMPONENTS METHOD OF FREQUENCY-RESPONSE MEASUREMENT

As in other measurements, null methods are generally superior to other methods in the matter of accuracy. More accurate measurements are possible with the phase-null method at the expense of somewhat greater measurement time.

Phase-Null Method. The method is shown in block-diagram form in Fig. 8. The sine and cosine-wave generator is used to provide a signal f_4 of adjustable phase α which is correlated with the process output in the correlation computer to give the modified cross correlation $\bar{\phi}_{42}$.

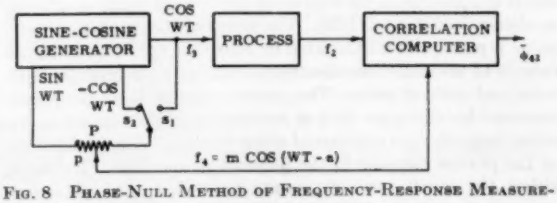


FIG. 8 PHASE-NULL METHOD OF FREQUENCY-RESPONSE MEASUREMENT

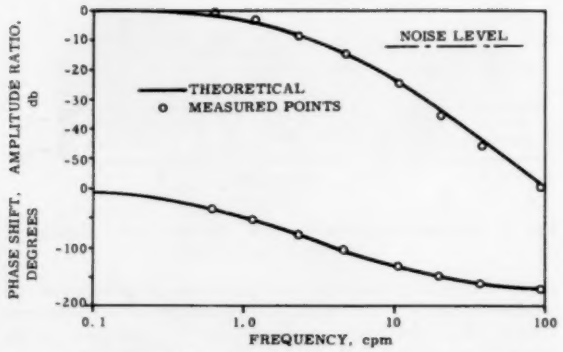


FIG. 9 AMPLITUDE RATIO AND PHASE SHIFT BY PHASE-NULL METHOD OF FREQUENCY-RESPONSE MEASUREMENT

The method of making measurements is described briefly in the following but is given more completely in the Appendix. Also given in the Appendix are details of the sine/cosine generator and the correlation computer.

Briefly the measurement procedure is as follows: (1) The potentiometer P is adjusted until the modified cross correlation ϕ_{42} is zero using switch position S_1 or S_2 as necessary. Let this potentiometer ratio be p_a. (2) The potentiometer is set to the complementary position and the switch S to the alternative position, so that the signal f_4 is 90 deg displaced in phase as compared with the phase resulting in the null balance. Let this value of the potentiometer ratio be p_b. Observe the value of the modified cross correlation ϕ_{42}.

The frequency response of the process has an amplitude ratio A and phase angle θ given by

$$A = \frac{2}{m_b}\,\bar{\phi}_{42} \qquad [3]$$

$$\theta = -\tan^{-1}\left(\frac{1-p_b}{p_b}\right), \text{ for } S \text{ at } S_1 \qquad [4a]$$

$$= -\pi + \tan^{-1}\left(\frac{1-p_b}{p_b}\right), \text{ for } S \text{ at } S_2 \qquad [4b]$$

where m_b is a correction coefficient which (assuming negligible loading of the potentiometer P) is given by

$$m_b = \sqrt{(1 - 2p_b + 2p_b^2)} \qquad [5]$$

Equations [4] and [5] are plotted in Figs. 10 and 11, respectively. The ambiguity in phase resulting from the inverse trigonometrical functions in Equation [4] is unimportant because it is possible to follow changes in phase angle with frequency from tests at low frequency where ambiguity may be removed by direct observation.

A plot of the frequency response measured by this method is given in Fig. 9 together with the theoretical curves. It will be seen that measurements are possible up to a frequency of 95 cpm with slightly better accuracy than those obtained by the quadrature-components method. It also will be observed that measurements are possible in the face of signal-to-noise ratios as adverse as almost −50 decibels (db). The power of the correlation technique is perhaps best illustrated by chart records. Fig. 12 shows records of the sine-wave disturbance and process response with noise and without noise. The process response is so completely obscured by the noise that it is quite impossible to obtain by visual inspection an estimate of either the amplitude or the phase of the process response in the presence of the noise. However, either the quadrature-components method or the phase-null

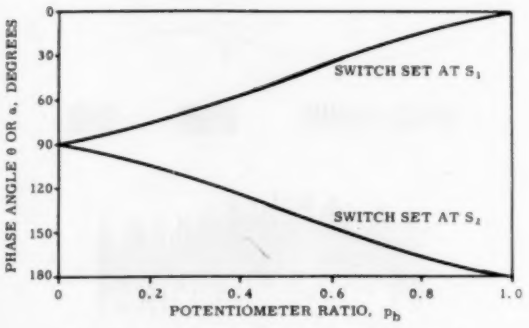


FIG. 10 PHASE ANGLE OF PROCESS OR SIGNAL f_4

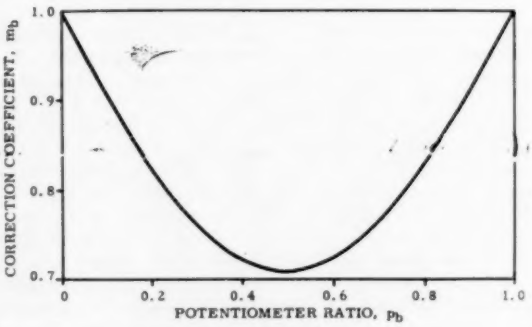


FIG. 11 CORRECTION COEFFICIENT, m_b

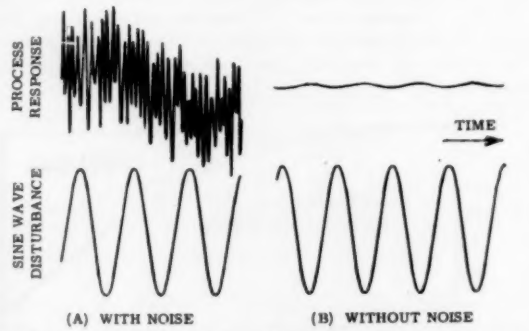


FIG. 12 CHART RECORDS OF SINE-WAVE DISTURBANCE AND PROCESS RESPONSE

method gives an accurate measure of both amplitude and phase.

EXPERIMENTAL RESULTS ON A REAL PROCESS

Measurement of Frequency Response on Closed Loop. In making a frequency-response measurement on a process it is desirable to remove the automatic control and to apply the sinusoidal disturbance to the process at the control valve. However, some processes cannot be taken "off control" (i.e., put onto "manual") for more than a few minutes at a time because of their natural instability. Many processes involving an integration in their transfer function (such as level control) may be taken off control for extended periods of time provided that occasional manual adjustments are made to maintain a zero average deviation of the process variable. There are some processes which when taken off control cannot be "steered" manually. Such processes

might involve in their transfer function either an integration with a large gain (such that the rate of integration is too fast for operator control) or a right-hand plane pole where there is a natural instability tending to cause the process variable to "run away."

Exothermic-temperature processes are generally of this type. Although the process cannot be taken off control, frequency-response measurements may still be made. The combination of process and controller is excited by applying the sinusoidal disturbance to either the controller set point or to the valve via a pneumatic adding relay, Fig. 13. The process frequency response is then obtained from the complex ratio C/P.

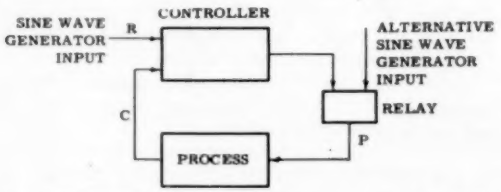


FIG. 13 FREQUENCY-RESPONSE MEASUREMENTS OF A PROCESS ON CLOSED LOOP

In employing the current method of frequency-response measurement the quantities C and P would be recorded on a chart and the amplitude ratio and phase shift would be obtained by visual reduction as in Fig. 1. In employing the correlation technique, the process input P, and the process output C, would be separately (and simultaneously) correlated with the sine-cosine generator giving the amplitude ratios A_1 and A_2 and phase shifts θ_1 and θ_2, respectively. The amplitude ratio A and phase shift θ of the process are then

$$A = \frac{A_2}{A_1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [6a]$$

$$\theta = \theta_2 - \theta_1 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [6b]$$

The difficulty encountered in measuring the frequency response of a process on closed loop arises from the fact that process noise restricts low-frequency measurements as well as high-frequency measurements. In contrast, open loop-frequency response measurements are restricted by noise only at high frequencies. Consequently process noise is a much more serious problem in the measurement of the process-frequency response on closed loop than in the measurement of the process-frequency response on open loop.

Frequency-response measurements on an exothermic temperature process measured by the visual method are given in Fig. 14. Measurements were attempted at octave intervals of frequency but useful results were obtained over a range of only two octaves.

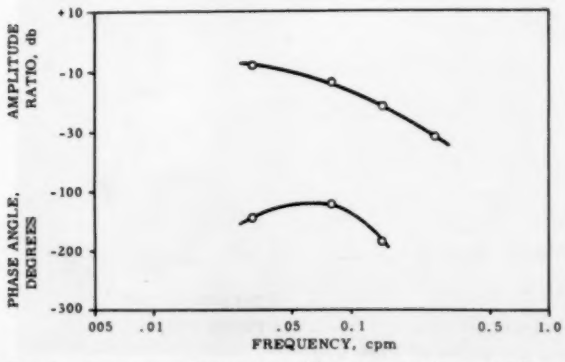


FIG. 14 FREQUENCY RESPONSE OF TEMPERATURE PROCESS MEASURED BY VISUAL METHOD

Figs. 15, 16, and 17 show the recorded wave forms of the sine-wave disturbance, the process input, and the process output at $^1/_2$, $^1/_8$, and $^1/_{32}$ cpm. Although the total frequency range is only 4 octaves it may be seen readily that amplitude ratio and the phase shift cannot be obtained from the wave forms excepting at the center frequency. At the "high" frequency ($^1/_2$ cpm) the process *output* has a poor signal-to-noise ratio[7] and at the low frequency ($^1/_{32}$ cpm) the process *input* has a poor signal-to-noise ratio. Consequently, the frequency range over which frequency-response measurements may be made by the visual method is severely limited.

The results of frequency-response measurements made by the correlation technique[8] are given in Fig. 18. Useful measurements were possible over a frequency range of about two decades or about 32 times the frequency range obtainable by the visual method. This comparison clearly demonstrates the advantage of computer reduction of data by the correlation method.

A photograph of the equipment setup for these measurements is shown in Fig. 19.

DISCUSSION OF RESULTS

Advantages of Correlation Technique. The foregoing results show that methods of frequency-response measurement using the correlation technique offer a substantial improvement in frequency-response measurement of "difficult" processes as compared with the visual method. Under the description of difficult processes we may list:

1 Processes into which only very small amplitude disturbances may be introduced.

2 Processes which are noisy in the sense that in the absence of control the process variable varies about its mean value in a random manner.

3 Processes which are subject to disturbances which, in the absence of control, cause the process variable to vary widely and thus cloak the response to a sine-wave disturbance introduced for measurement purposes.

4 Inherently unstable systems which in the absence of control would run away. Included in these are the temperature processes of exothermic reactions. Even for processes which are not difficult in any of the senses mentioned, correlation techniques offer a means of extending the upper frequency limit to which measurements may be made. This is often desirable for systems for which rapid control is required, so that the value of rate action may be assessed properly.

The correlation technique also enables very accurate frequency-response measurements to be obtained. This is an important factor in the measurement and analysis of systems comprising coupled processes.

Disadvantages of Correlation Techniques. Among the disadvantages we may list:

1 The requirement for an analog computer containing at least one but preferably two accurate and stable multipliers and integrators.

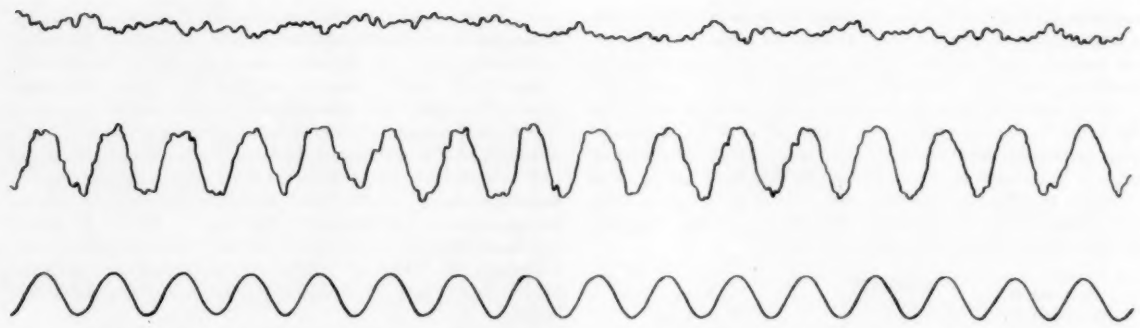
2 The need to have the computer located at the test site.

3 The time required to make the measurements.

It is not necessary that an electronic analog computer be used although the experimental work was carried out with one. Mechanical ball-and-disk integrators may be employed for both

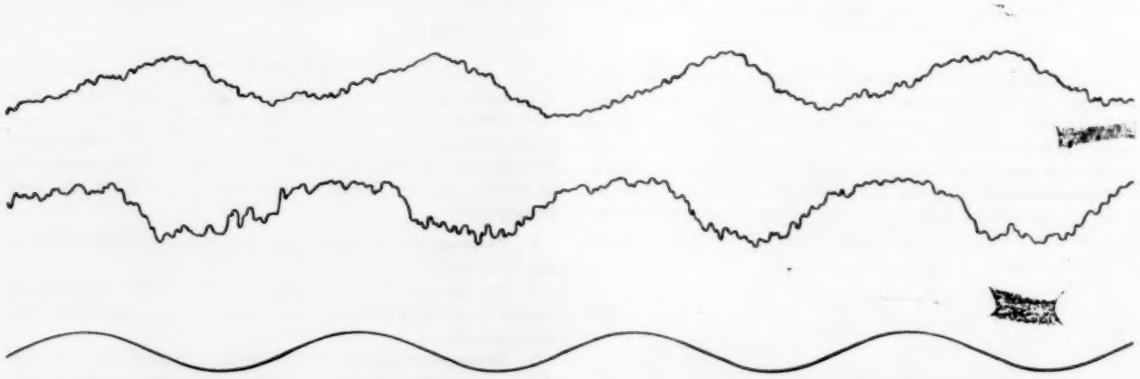
[7] Signal here means the amplitude of the sine wave component having the same frequency as the disturbing sine wave.

[8] Using the Quadrature Components Method.



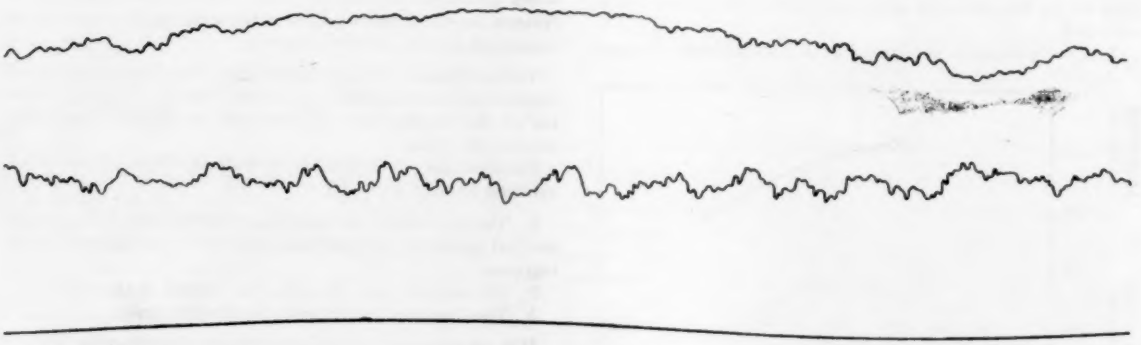
From Top to Bottom: Temperature — Process Output
Heat-Medium Valve (air pressure) — Process Input
Sine-Wave Generator (at set point of T. R. C.)

FIG. 15 WAVE FORMS OF SINE-WAVE DISTURBANCE, PROCESS INPUT, AND PROCESS OUTPUT; HIGH FREQUENCY, 1/2 CPM



From Top to Bottom: Temperature — Process Output
Heat-Medium Valve (air pressure) — Process Input
Sine-Wave Generator (at set point of T. R. C.)

FIG. 16 WAVE FORMS OF SINE-WAVE DISTURBANCE, PROCESS INPUT, AND PROCESS OUTPUT; MEDIUM FREQUENCY, 1/8 CPM



From Top to Bottom: Temperature — Process Output
Heat-Medium Valve (air pressure) — Process Input
Sine-Wave Generator (at set point of T. R. C.)

FIG. 17 WAVE FORMS OF SINE-WAVE DISTURBANCE, PROCESS INPUT, AND PROCESS OUTPUT; LOW FREQUENCY, 1/32 CPM

multiplier and integrator in a correlation computer. Alternatively, multiplication may be carried out in the transducer with which the controlled variable is measured, but the stability requirements indicate that mechanical integration would be required.

The time to obtain each measurement by the correlation technique is much greater than the time required for measurement by the visual method. Each observation must be made over a large number of cycles (perhaps 15 or 20) of the disturbing sine wave, and several observations may be necessary to obtain a null balance. However, the situation is not as serious as appears at first sight because at the high frequencies (where the correlation technique is of greatest[9] value) the periods are short and the additional time requirements are not excessive.

At the lower frequencies, measurements may be made by the visual method in order to conserve time. Alternatively, since noise and distortion may reasonably be expected to be a small fraction of the total system response, the correlation technique may be employed using integration over a small number of cycles.

Limitation of Analog Computation

The chief limitation in analog computation of the correlations lies in the limited stability of the integrators and accuracy of the multipliers. These limitations may be removed by resorting to numerical computation.

Numerical Computation. Provided that the variables can be digitized and recorded accurately, the computation of the correlations is readily performed with a digital computer. Integrator drift and multiplier inaccuracies are virtually nonexistent.

[9] This refers to frequency-response measurements of processes on open loop. For frequency-response measurements of processes on closed loop the correlation technique is of equal value at low frequencies and observation over a large number of cycles involves a somewhat excessive time.

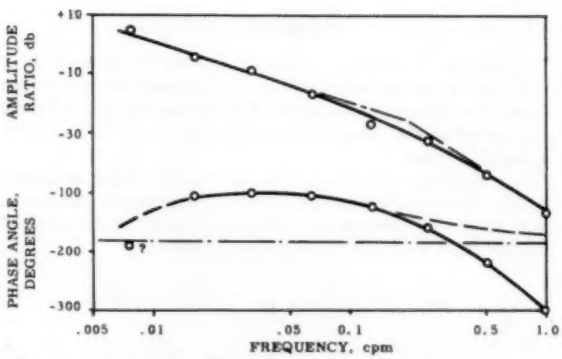


Fig. 18 Frequency Response of Temperature Process Measured by Correlation Method

Similarly the provision of a delay for computation of the correlation function proper is much less a problem than in the case of analog computation. The major problem is that of digitizing and recording the variables. The sampling rate should be at least twice, and preferably ten times the frequency of the disturbing sine wave.

Conclusion

Noise and distortion may seriously limit the frequency range over which frequency-response measurements can be made by the visual method. In such cases the correlation methods described in this paper enable measurements to be made over a much wider frequency range. They are particularly useful for the measurement of the frequency response of processes (such as exothermic-temperature processes) which cannot be taken off control.

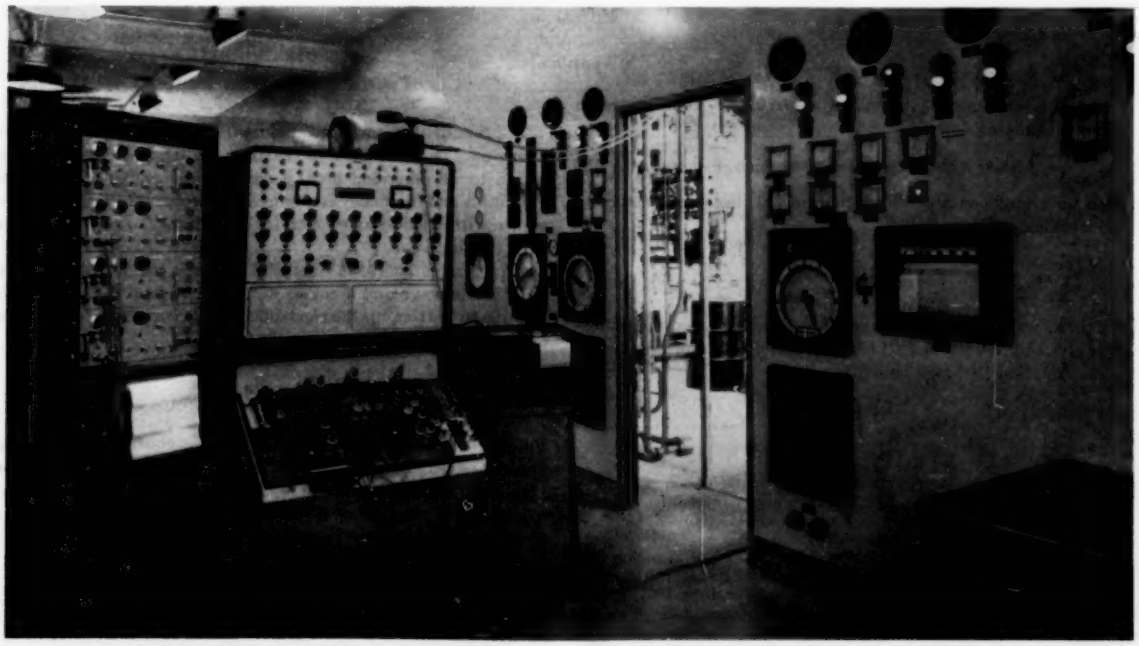


Fig. 19 Equipment Setup for Frequency-Response Measurement of Temperature Process

Appendix

Correlation Technique. Several papers[10-12] recently have appeared on the subject of extraction of a sine-wave signal from noise. The theory is not new, but it is believed that the following application of the theory to the measurement of frequency response is original:

In computing correlation functions with the analog computer, a delay or dead time of variable delay τ would permit the evaluation of the auto-correlation function $\phi_{11}(\tau)$ and the cross-correlation functions $\phi_{12}(\tau)$, $\phi_{21}(\tau)$ given by the following

$$\phi_{11}(\tau) = \underset{T \to \infty}{\text{Limit}} \frac{1}{2T} \int_{-T}^{+T} f_1(t) f_1(t+\tau) dt \quad [7]$$

$$\phi_{12}(\tau) = \underset{T \to \infty}{\text{Limit}} \frac{1}{2T} \int_{-T}^{+T} f_1(t) f_2(t+\tau) dt \quad [8a]$$

$$\phi_{21}(\tau) = \underset{T \to \infty}{\text{Limit}} \frac{1}{2T} \int_{-T}^{+T} f_2(t) f_1(t+\tau) dt \quad [8b]$$

Since pure delay or dead time is very difficult to achieve for analog computation, the expressions given in Equations [7], [8*a*], and [8*b*] are difficult to evaluate even approximately. However, the cross correlation for zero delay is extremely useful in certain circumstances

$$\phi_{12}(0) = \phi_{21}(0) = \underset{T \to \infty}{\text{Limit}} \frac{1}{2T} \int_{-T}^{T} f_1(t) f_2(t) dt \quad [9]$$

A useful approximation to the expression in Equation [9] can be obtained by eliminating consideration of future time and not proceeding to the limit indicated. The "zero delay" correlation is then a function of the integration period T and time t

$$\phi_{12}(T, t) = \frac{1}{T} \int_{-T}^{0} f_1(t) f_2(t) dt \quad [10]$$

If $f_1(t)$ is a sine wave

$$f_1(t) = \sin \omega t \quad [11]$$

and $f_2(t)$ has a component of the same frequency plus distortion terms plus noise

$$f_2(t) = a \sin(\omega t - \theta) + \text{distortion} + \text{noise} \quad [12]$$

then the expression given by Equation [10] becomes practically independent of T for

$$T \gg \frac{2\pi}{\omega} \quad [13]$$

Furthermore if $T = 2\pi N/\omega$, where N is a large integer, the expression given by Equation [10] becomes practically independent of t and Equation [10] may be written

$$\phi_{12} = \frac{1}{T} \int_{-T}^{0} f_1(t) f_2(t) dt \quad [14]$$

A more convenient quantity to compute on the analog computer is the running integral

$$\bar{\phi}_{12}(t) = \frac{1}{T_0} \int_{0}^{t} e^{-(t-x)/T_0} f_1(x) f_2(x) dx \quad [15]$$

If the time constant T_0 is very much greater than the period $2\pi/\omega$ of the sine wave, then the expression in Equation [15] becomes practically independent of t for $t > 5T_0$. The quantity $\bar{\phi}_{12}$ is a measure of the sine-wave component of frequency ω in $f_2(t)$. Similarly, if f_1 is replaced by $f_3 = \cos \omega t$, then $\bar{\phi}_{32}$ becomes a measure of the cosine-wave component in $f_2(t)$.

Practical Circuitry

The Correlation Computer. The modified cross correlations $\bar{\phi}_{12}$, $\bar{\phi}_{32}$, and $\bar{\phi}_{42}$ may be computed with the electronic analog computer as shown in the block diagram of Fig. 20. The use of a potentiometer in the resistance feedback path of the operational amplifiers enables very large time constants to be realized with readily obtainable components. Normally the time constants T_1 and T_2 would be made equal.

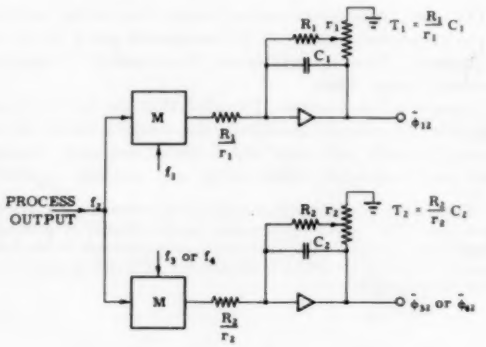


Fig. 20 Correlation Computer

The Sine-Wave Generator. Before frequency-response measurements may be made a sine-wave source is required. In the application to be described here, a cosine-wave source of the same amplitude and frequency is required. It would be possible to commence with a sine-wave generator and derive[13] the cosine wave from the sine wave.

A preferable course is to set up a circuit which will generate both a sine wave and a cosine wave simultaneously. This may be done by setting up the electronic analog computer to solve the harmonic equation

$$\frac{d^2x}{dt^2} + \omega^2 x = 0 \quad [16]$$

Practical experience indicates that there is considerable value in applying a little negative damping to the circuit and limiting the amplitude of the oscillation build-up. The wave form is not distorted provided that the limiting is accomplished with a diode and bias of low impedance. Equation [16] may be replaced by

$$\frac{d^2x}{dt^2} - 2k\omega \frac{dx}{dt} + \omega^2 x = 0, \quad k \ll 1 \quad [17]$$

[10] "Application of Correlation Analysis to the Detection of Periodic Signals in Noise," by Y. W. Lee, T. P. Cheatham, and J. B. Weisner, Proceedings of the Institution of Radio Engineers, vol. 38, October, 1950, pp. 1165–1171.

[11] "Perturbation and Correlation Methods for Enhancing the Space Resolution of Directional Receivers," by F. V. Hunt, Proceedings of the Institution of Radio Engineers, vol. 39, July, 1951, p. 840.

[12] "The Detection of a Sine Wave in the Presence of Noise by the Use of a Nonlinear Filter," by T. G. Slattery, Proceedings of the Institution of Radio Engineers, vol. 40, October, 1952, pp. 1232–1236.

[13] Such a circuit would need adjustment for each frequency.

$$x_{max} \gg E_b \quad \text{[18]}$$

The analog-computer diagram for generating sine and cosine waves by solving Equations [17] and [18] is given in Fig. 21.

Frequency-Response Measurements

Quadrature-Components Method. The quadrature-components method of frequency-response measurement using the correlation technique is shown in Fig. 6. In this figure the blocks representing the correlation computer and the sine-and-cosine-wave generator are as shown in Figs. 20 and 21, respectively.

The transfer function of the process has an amplitude ratio A and phase angle θ which may be determined from $\bar{\phi}_{12}$ and $\bar{\phi}_{22}$

$$A = 2\sqrt{[(\bar{\phi}_{12})^2 + (\bar{\phi}_{22})^2]} \quad \text{[19]}$$

$$\theta = -\tan^{-1}\frac{\bar{\phi}_{12}}{\bar{\phi}_{22}} \quad \text{[20]}$$

The quantities $\bar{\phi}_{12}$ and either of $\bar{\phi}_{22}$ or $\bar{\phi}_{42}$ require a computation time of about five time constants. They are obtained readily by recording the quantities on a chart and reading off the average steady-state values. Fig. 22 is a typical record of the development of these cross correlations.

If in the circuit of Fig. 20 the operational amplifiers are used as pure integrators, the mean slope of the record gives the value of the modified cross correlations. Fig. 23 shows a typical record when following this alternative procedure. With very low-frequency sine-wave disturbances the record contains a double frequency component as shown in Fig. 24. By marking the $^1/_2$ periods on the record and drawing two best straight lines through these points, the cross correlations are obtained from the slopes of these lines.

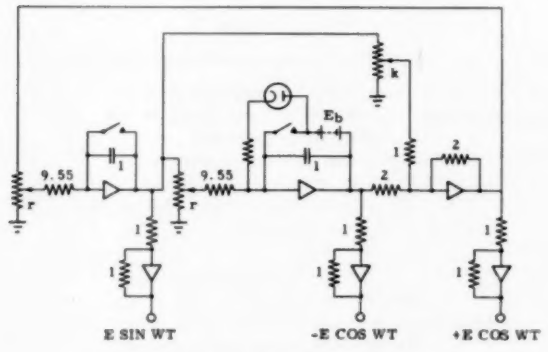


Fig. 21 Sine-Cosine Generator

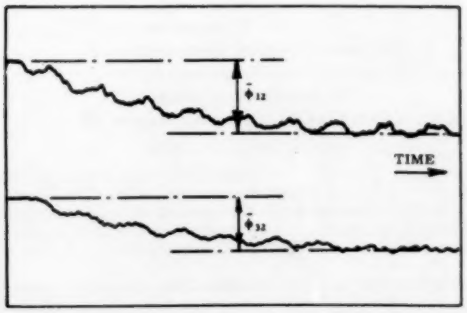


Fig. 22 Typical Chart Record Obtained With Quadrature-Components Method

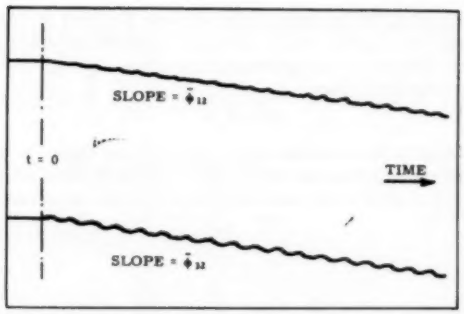


Fig. 23 Typical Chart Record Obtained With Quadrature-Components Method Using Alternative Procedure

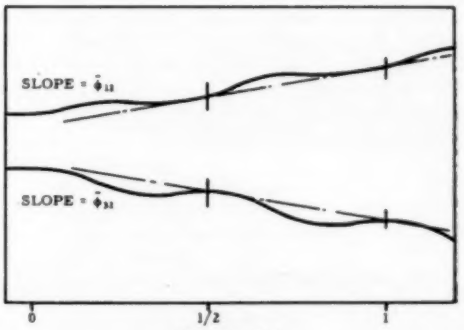


Fig. 24 Chart Record Obtained With Quadrature-Components Methods Showing Double-Frequency Component

Phase-Null Method. The phase-null method of frequency-response measurement using the correlation technique is shown in Fig. 8. In this figure the blocks representing the correlation computer and the sine-and-cosine generator are as shown in more detail in Figs. 20 and 21, respectively.

The phase of the generator signal f_4 with respect to f_3 is given by

$$\alpha = -\tan^{-1}\left(\frac{1-p}{p}\right) \text{ for } S \text{ at } S_1 \quad \text{[21]}$$

or

$$\alpha = -\pi + \tan^{-1}\left(\frac{1-p}{p}\right) \text{ for } S \text{ at } S_2 \quad \text{[22]}$$

These equations are plotted in Fig. 10. The amplitude of the signal f_4 varies between 0.707 and 1.000 times the amplitude of the signal f_3 as the potentiometer ratio is varied. If the loading on the potentiometer P is negligible the correction coefficient is given by

$$m_b = \sqrt{(1 - 2p_b + 2p_b^2)} \quad \text{[23]}$$

and is plotted in Fig. 11.

The procedure used to obtain the frequency response is as follows:

1 Set the generator to oscillate at the frequency at which the response is desired.

2 Adjust the potentiometer P until the modified cross correlation $\bar{\phi}_{42}$ is zero using switch position S_1 or S_2 as necessary. Let this potentiometer ratio be p_a.

3 Set the phase of the signal f_4 to be 90 deg displaced from the

phase resulting in the null balance in item 2. This is done by throwing the switch S to the alternative position and setting the potentiometer to the complementary position. Let this potentiometer ratio be p_b, then

$$p_a + p_b = 1 \quad [24]$$

4 Observe the modified cross correlation $\bar{\phi}_{42}$ with $p = p_b$.

The frequency response of the process has an amplitude ratio A and phase angle θ which may be determined from p_b, m_b, and $\bar{\phi}_{42}$

$$A = \frac{2}{m_b} \bar{\phi}_{42} \quad [25]$$

$$\theta = -\tan^{-1}\left(\frac{1 - p_b}{p_b}\right)$$

The ambiguity in the phase angle is unimportant because it is possible to follow changes in phase angle with frequency from tests at low frequency where ambiguity may be removed by direct observation. By duplicating part of the equipment shown in Fig. 8, it is possible to make the two computations involved in steps 2 and 4 concurrently instead of sequentially. Fig. 25 shows a chart record in which the null test and the modified cross-correlation function $\bar{\phi}_{42}$ are obtained concurrently.

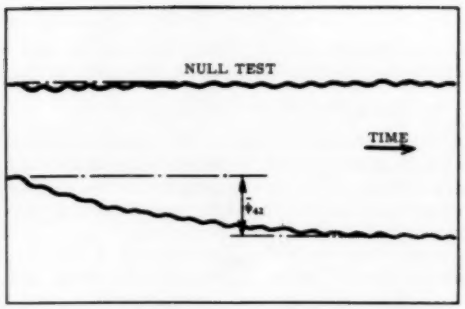


FIG. 25 TYPICAL CHART RECORD OBTAINED WITH PHASE-NULL METHOD

If in the circuit of Fig. 20 the operational amplifiers are used as pure integrators, the lower curve in Fig. 25 no longer will be exponential but will have a mean slope which is proportional to the modified cross correlation $\bar{\phi}_{42}$. The mean slope is best obtained by marking the half-period points on the record and drawing the best straight line through them.

Criteria for Validity of Lumped-Parameter Representation of Ducting Air-Flow Characteristics

By T. R. STALZER[1] AND G. J. FIEDLER,[2] ST. LOUIS, MO.

The dynamic characteristics of air flow in a section of ducting involve nonlinear partial differential equations in which the air pressure and weight flow vary with both space and time. Differential equations for the distributed-parameter and the lumped-parameter representations are derived. Using applicable linearizing assumptions, transfer functions are obtained for each representation. The two transfer functions are compared by applying them to a practical valve-and-ducting system selected from a control problem. Since the lumped-parameter representation is more convenient to use in controls design, a simple numerical criterion is developed to determine rapidly the conditions under which lumped representation of a distributed element is valid. Frequency-response and time-response plots for the lumped and distributed elements are included for comparison purposes.

Nomenclature

The following nomenclature is used in the paper:

$a_1 = A_1^* + A_1$ = instantaneous value of valve 1 open area, sq ft
A_1^* = fixed, base, or steady-state value of a_1 at an operating point, sq ft
A_1 = incremental variation in a_1 about A_1^*, sq ft
A_d = cross-sectional area of the ducting, sq ft
$\bar{F}_1, \bar{F}_1^*, F_1(s)$ = valve 1 function of p_2/p_1
g = acceleration of gravity = 32.2 ft/sec²
j = $(-1)^{1/2}$
M = weight of air, lb
$p = P^* + P$ = air pressure, psfa
$w = W^* + W$ = weight flow of air, lb/sec
R = gas constant = 53.3 ft/deg R
s = Laplace-transform complex variable with respect to time t, sec^{-1}
t = time, sec
u = velocity of air movement, fps
u_c = velocity of sound in air, or acoustic velocity, fps
V = volume, cu ft
x = distance or length along ducting, ft
z = Laplace transform complex variable with respect to distance x, ft^{-1}
Z_0 = characteristic or surge impedance of the ducting
$K_c = \frac{gA_d}{\mu_c^2}$, pneumatic capacitance per unit length
$K_n = \frac{1}{gA_d}$, pneumatic inertance per unit length
$\alpha = s/u_c = s\sqrt{(K_c K_n)}$
$\beta = sx/u_c$ = ducting propagation constant
$\gamma = c_p/c_v = 1.40$ (experimental value)
$\theta = \Theta^* + \Theta$ = air temperature, deg R
λ = wave length corresponding to a frequency ω and velocity of sound u_c for air flow in ducting
$\mu(t)$ = unit step function
$\pi = 3.1416$
ρ = weight density of air, pcf
$\sigma = 0.0282s$
ω = angular frequency, radians/sec

Introduction

The flow of compressible fluid in a long pipe or duct is somewhat analogous to the flow of electric current in a long transmission line. For accurate representation of gaseous flow at high frequencies in a long duct, the resistance, inertance, and capacitance parameters must be considered as distributed along the duct. Air pressures and weight flow are functions of time and distance along the duct.

For control-design purposes, the lumped-parameter representation of ducting elements is more convenient and simpler to use. The purpose of this paper is to develop a simple general numerical criterion to determine rapidly the conditions under which lumped representation of a distributed ducting element is valid. To develop this criterion, equations and transfer functions for the distributed-parameter as well as the lumped-parameter representation are derived and compared for a practical ducting-and-valve system.

DISTRIBUTED-PARAMETER DUCTING AIR-FLOW CHARACTERISTICS

In the analysis of fluid motion, it is not necessary to treat the fluid as being composed of discrete particles. It is permissible to consider the fluid as being composed of elementary volumes as shown in Fig. 1. In this development, the partial differential

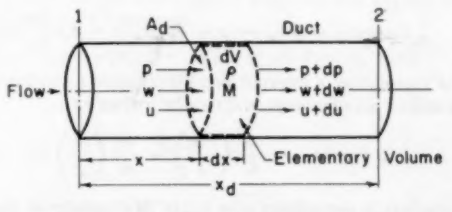


Fig. 1 Elementary Fluid Volume in a Straight Duct

[1] McDonnell Aircraft Corporation.
[2] Principal Engineer, Sverdrup & Parcel, Inc.

Contributed by the Instruments and Regulators Division of The American Society of Mechanical Engineers and presented at the joint ASME-ISA Conference, New York, N. Y., September 17–21, 1956.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, July 3, 1956. Paper No. 56—IRD-21.

equations relating the flow and pressure of a fluid with time, distance, and ducting parameters will be determined. Two equations will be utilized; namely, the continuity equation and the force equation.

The continuity equation (1)[2] states that the rate of change of mass within the elementary volume, due to the changing density, is equal to the net mass flow into the element. From Fig. 1, the rate of change of mass within the elementary fixed volume may be expressed as

$$\frac{\partial M}{\partial t} = w - (w + dw) = -dw = -\frac{\partial w}{\partial x}\,dx \quad [1]$$

where

M = weight of air within element
w = weight flow of air into element

However, the time rate of change of mass also can be written as

$$\frac{\partial M}{\partial t} = A_d \frac{\partial \rho}{\partial t}\,dx \quad [2]$$

where

A_d = flow cross-sectional area
ρ = weight density of air
dx = elementary length of element

For an isentropic process (1)

$$p\rho^{-\gamma} = \text{const} \quad \text{and} \quad \frac{\gamma p}{\rho} = \frac{u_c^2}{g}$$

therefore

$$\frac{\partial p}{\partial t} = \frac{g}{u_c^2}\frac{\partial p}{\partial t} \quad [3]$$

where

γ = ratio of specific heats of air
u_c = acoustic velocity
g = gravitational constant
p = air pressure

Combining Equations [1], [2], and [3] gives

$$\left(\frac{gA_d}{u_c^2}\right)\frac{\partial p}{\partial t} = -\frac{\partial w}{\partial x} \quad [4]$$

The force equation (1) states that the resultant external force acting on the elementary volume is equal to the time rate of change of the resultant momentum within the element, plus the net outflow of momentum through the surface of the elementary volume. For simplification in writing the force equations, friction and leakage effects are omitted. It is noted that wall friction affects the velocity profile and hence the inductance parameter; however, these effects are negligible in the example treated. Furthermore, the inclusion of frictional forces would violate the isentropic relation utilized in Equation [3]. Referring again to Fig. 1, the summation of external forces may be expressed as

$$\Sigma \text{ Forces}_{ext} = pA_d - (p + dp)A_d = -A_d\frac{\partial p}{\partial x}\,dx \quad [5]$$

The summation of internal forces is obtained by utilizing Newton's second law of motion, and may be written as

$$\Sigma \text{ Forces}_{int} = \frac{d}{dt}\left(\frac{Mu}{g}\right) = \frac{d}{dt}\left(\frac{w}{g}\right)dx \quad [6]$$

since

$$M = \rho A_d\,dx \quad \text{and} \quad w = \rho A_d u \quad [7]$$

and u is the flow velocity.

Equation [6] may be expanded as

$$\Sigma \text{ Forces}_{int} = \frac{1}{g}\frac{\partial w}{\partial t}\,dx + \frac{1}{g}\frac{\partial w}{\partial x}\frac{dx}{dt}\,dx = \frac{1}{g}\frac{\partial w}{\partial t}\,dx + \frac{u}{g}\frac{\partial w}{\partial x}\,dx \quad [8]$$

(where $dx/dt = u$, the flow velocity) since x is defined as the position of the elementary volume along the ducting.

Combining Equations [5] and [8] produces

$$-gA_d\frac{\partial p}{\partial x} = \frac{\partial w}{\partial t} + u\frac{\partial w}{\partial x} \quad [9]$$

Utilizing Equations [3], [4], and [7], Equation [9] can be written as

$$-gA_d\frac{\partial p}{\partial x} = \frac{\partial w}{\partial t} - \frac{w}{\gamma p}\frac{\partial p}{\partial t} \quad [10]$$

Assuming typical values for w and p as given later in the paper, $w/\gamma p = 0.028$. Thus, for engineering accuracy the contribution of the $(w/\gamma p)(\partial p/\partial t)$ term is negligible compared to the sum of the other two terms in Equation [10]. Since this term of Equation [10] is assumed to be negligible, Equation [9] may be simplified to the linear form

$$\left(\frac{1}{gA_d}\right)\frac{\partial w}{\partial t} = -\frac{\partial p}{\partial x} \quad [11]$$

It is appropriate to assume small perturbation conditions for the variables w and p for reasons which will be apparent later.

For small perturbations, $w = W^* + W$ and $p = P^* + P$, where W and P are incremental variables. Then Equations [4] and [11] become

$$K_c\frac{\partial P}{\partial t}(x, t) = -\frac{\partial W}{\partial x}(x, t) \quad [12]$$

$$K_n\frac{\partial W}{\partial t}(x, t) = -\frac{\partial P}{\partial x}(x, t) \quad [13]$$

where

$K_c = gA_d/u_c^2$, pneumatic capacitance/unit length

$K_n = 1/gA_d$, pneumatic inertance/unit length

Equations [12] and [13] are now analogous to the equations for voltage and current as functions of time and space along an electric transmission line having zero conductance and resistance (2). To assume either resistance (loss of pressure) or conductance (loss of flow due to leakage), would violate the isentropic assumption. Furthermore, it can be shown that the duct resistance is small compared to the resistance of components such as valves, compressors, and so on.

In order to obtain the transfer function for the ducting element, it is necessary to derive the expressions for $P(x, s)$ and $W(x, s)$, where s is the Laplace-transform variable with respect to time. The Laplace transforms (3) of Equations [12] and [13] may be written as

$$sK_cP(x, s) = -\frac{dW}{dx}(x, s) \quad [14]$$

$$sK_nW(x, s) = -\frac{dP}{dx}(x, s) \quad [15]$$

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

Combining Equations [14] and [15] gives

$$\frac{d^2P}{dx^2}(x, s) - \alpha^2 P(x, s) = 0 \quad [16]$$

and

$$\frac{d^2W}{dx^2}(x, s) - \alpha^2 W(x, s) = 0 \quad [17]$$

where

$$\alpha^2 = K_c K_n s^2 = s^2/u_e^2$$

Writing the Laplace transform with respect to x, Equations [16] and [17] become

$$z^2 P(z, s) - zP(0, s) - \frac{dP}{dx}(0, s) - \alpha^2 P(z, s) = 0 \quad [18]$$

$$z^2 W(z, s) - zW(0, s) - \frac{dW}{dx}(0, s) - \alpha^2 W(z, s) = 0 \quad [19]$$

where z is the Laplace variable with respect to x.

Combining Equations [14], [15], [18], and [19] gives

$$P(z, s) = \frac{zP(0, s)}{z^2 - \alpha^2} - \frac{sK_n W(0, s)}{z^2 - \alpha^2} \quad [20]$$

and

$$W(z, s) = \frac{zW(0, s)}{z^2 - \alpha^2} - \frac{sK_c P(0, s)}{z^2 - \alpha^2} \quad [21]$$

The inverse Laplace transforms (3) of Equations [20] and [21] with respect to z may be written as

$$P(x, s) = P(0, s)\cosh\beta - \frac{s}{\alpha} K_n W(0, s)\sinh\beta \quad [22]$$

$$W(x, s) = W(0, s)\cosh\beta - \frac{s}{\alpha} K_c P(0, s)\sinh\beta \quad [23]$$

where $\beta = \alpha x = sx/u_e$, the propagation constant. The surge impedance Z_0 (2, 4) is defined as

$$Z_0 = \sqrt{(K_n/K_c)} = sK_n/\alpha = \alpha/sK_c \quad [24]$$

The downstream-end impedance is

$$Z_2^* = P_2^*/W_2^* = P_2(s)/W_2(s) \quad [25]$$

where the asterisk indicates the use of fixed values.

Combining Equations [22], [23], and [24] produces

$$P_2(s) = P_1(s)\cosh\beta - Z_0 W_1(s)\sinh\beta \quad [26]$$

$$W_2(s) = W_1(s)\cosh\beta - \frac{1}{Z_0} P_1(s)\sinh\beta \quad [27]$$

for $x = x_d$. In Equations [26] and [27]

$$\beta = sx_d/u_e$$

The distributed-parameter transfer functions for ducting may be obtained from Equations [25], [26], and [27], which are analogous to the equations (2, 4) for a long electric transmission line having distributed parameters.

LUMPED-PARAMETER, DUCTING AIR-FLOW CHARACTERISTICS

The lumped-parameter representation of ducting characteristics is normally used for most applications. It is less complicated mathematically and yet provides a means for handling changes in pressure as well as temperature. The theory on which the mathematical derivation is based involves the assumptions that the process is isentropic and that the ducting is equivalent to a single lumped volume in which the pressure and temperature are constant and uniform throughout the volume. Although the distributed-parameter equations include the effects of pneumatic inertance and capacitance, the lumped-parameter equation includes only the effect of pneumatic capacitance. A lumped-parameter representation may include inertance as well as capacitance. However, for the example treated, a single resonance peak is present between the lumped-parameter break frequency and the first distributed-parameter peak on a plot similar to Fig. 3. Therefore, pneumatic capacitance only was used for the lumped-parameter representation.

Referring to Fig. 2, the rate of change of mass within the

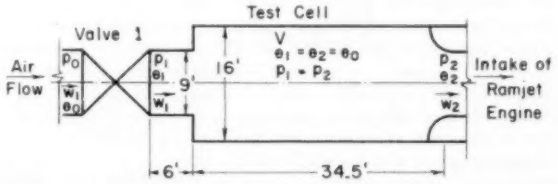


FIG. 2 TEST CELL AND CONTROL VALVE

lumped volume is equal to the net weight flow into the volume, and may be written as

$$w_1 - w_2 = \frac{dM}{dt} = \frac{V}{R\theta_1}\frac{dp_1}{dt} - \frac{Vp_1 d\theta_1}{R\theta_1^2 dt} \quad [28]$$

since $M = p_1 V/R\theta_1$. If the changes of state within the lumped volume occur isentropically

$$p_1(\theta_1)^{\gamma/1-\gamma} = \text{const} \quad [29]$$

The time derivative of Equation [29] is

$$\frac{p_1}{\theta_1}\frac{d\theta_1}{dt} = \frac{(\gamma - 1)}{\gamma}\frac{dp_1}{dt} \quad [30]$$

Combining Equations [28] and [30] produces

$$w_1 - w_2 = \frac{V}{\gamma R\theta_1}\frac{dp_1}{dt} \quad [31]$$

Since w_1, p_1, and θ all vary with time, it is necessary to use the small perturbation linearizing technique so that

$$w_1 = W_1^* + W_1$$

$$w_2 = W_2^* + W_2$$

$$\theta_1 = \Theta_1^* + \Theta_1 = \theta_2$$

$$p_1 = P_1^* + P_1 = p_2$$

where W_1, etc., are incremental variables.

Taking first the total differential and then the Laplace transform, Equation [31] becomes

$$dw_1 - dw_2 = W_1(s) - W_2(s) = sK_7 P_2(s) + K_8\Theta_1(s) \quad [32]$$

where

$$K_7 = \frac{V}{\gamma R\Theta_1^*}$$

$$K_8 = -\frac{V\dot{P}_2^*}{\gamma R(\Theta_1^*)^2}$$

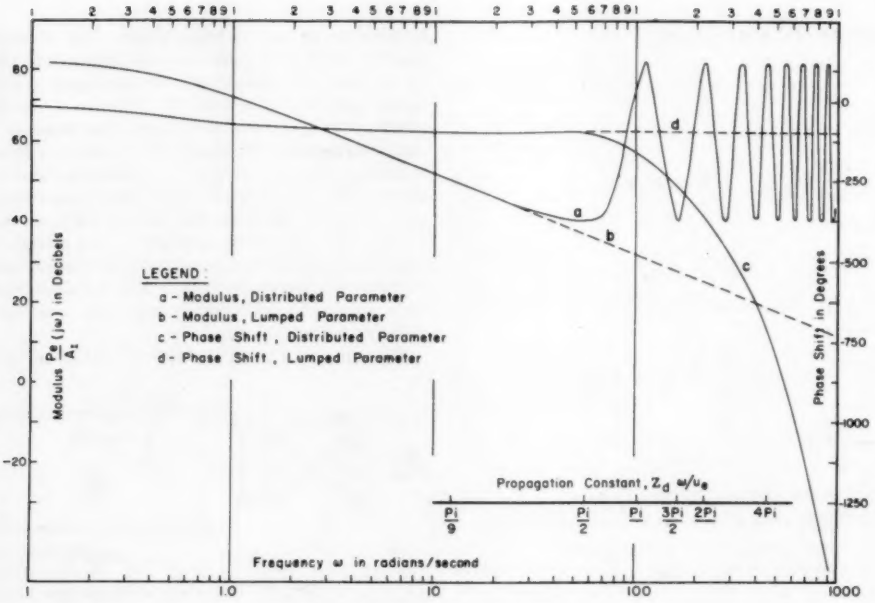


FIG. 3 DUCTING FREQUENCY RESPONSE

A CRITERION FOR DUCTING-PARAMETER SELECTION

The question arises as to when the lumped parameter representation suffices. All fluids exhibit some capacitance as well as inertance characteristics. Rigorous analysis must consider these effects in the fluid, as well as pipe-wall effects and terminating-load effects. In general for engineering design, pneumatic inertance may be neglected for a compressible fluid, and pneumatic capacitance may be neglected for an incompressible fluid. Therefore, the lumped-parameter representation is considered sufficient for any fluid, except for long lines or high frequencies. In order to demonstrate the applicability of the lumped or distributed-parameter equations, an example will be presented using design data from a ramjet-engine test facility.

Referring to Fig. 2, let the valve upstream conditions be p_0 and θ_0; the valve downstream conditions be w_1, p_1, and θ_1; and the engine-inlet conditions be w_2, p_2, and θ_2. Assuming $\theta_0 = \theta_1 = \theta_2$ (const), p_0 const, a choked valve, and no engine disturbance, the lumped-parameter equations are

$$W_1 - W_2 = sK_7P_2, \text{ from Equation [32]} \quad [33]$$

where

$$W_1 = K_1A_1, \text{ from Equation [55]} \quad [34]$$

and

$$W_2 = K_5P_2, \text{ from Equation [58]} \quad [35]$$

From Equations [33], [34], and [35], the transfer function P_2/A_1 may be written as

$$\frac{P_2}{A_1}(s) = \frac{P_1}{A_1}(s) = \frac{K_1/K_5}{1 + sK_7/K_5} \quad [36]$$

For a sinusoidal variation in A_1, Equation [36] becomes

$$\frac{P_2}{A_1}(j\omega) = \frac{K_1/K_5}{1 + j\omega K_7/K_5} \quad [37]$$

The distributed-parameter relations include Equations [26], [27], [34], and [35]. Combining these equations gives the transfer function

$$\frac{P_2}{A_1}(s) = \frac{K_1/K_5}{\cosh \beta + \frac{1}{K_5Z_0} \sinh \beta} \quad [38]$$

For a sinusoidal variation in A_1, Equation [38] becomes

$$\frac{P_2}{A_1}(j\omega) = \frac{K_1/K_5}{\cos \omega x_d/u_e + \frac{j}{K_5Z_0} \sin \omega x_d/u_e} \quad [39]$$

since $\beta = j\omega x_d/u_e$.

A comparison of Equations [37] and [39] reveals a definite similarity between the two expressions. For values of $x_d\omega/u_e$ less than $\pi/9$, Equation [39] simplifies to Equation [37], since $Z_0 = 1/K_5u_e$, and $x_dK_c = K_7$. In order to make a numerical comparison, the data for a typical operating point of the ramjet engine test facility will be utilized. These data are

$P_1^* = P_2^* = 10{,}000$ psfa

$W_1^* = W_2^* = 300$ lb/sec

$\Theta_1^* = \Theta_2^* = 900$ deg R

$A_1^* = 1.022$ sq ft

$A_d = 128.5$ sq ft

$x_d = 40.5$ ft

$V = A_dx_d = 5210$ cu ft

Utilizing these data, the system parameters are evaluated as follows:

$u_e = 1470$ fps

$Z_0 = 0.355$ sec^2/ft^2

$K_1 = 293$ lb/sec ft^2

$K_4 = 0.030$ ft³/sec

$K_7 = 0.0777$ sq ft

Utilizing these numerical data, Equations [37] and [39] become

$$\frac{P_2}{A_1}(j\omega) = \frac{9760}{1 + j2.59\omega} \quad [40]$$

$$\frac{P_2}{A_1}(j\omega) = \frac{9760}{\cos(0.0282\omega) + j93.9\sin(0.0282\omega)} \quad [41]$$

Frequency-response plots of Equations [40] and [41] are shown in Fig. 3. It is seen that there is very good agreement (5) between curves *a* and *b* out to $\omega x_d/u_c = \pi/9$. For $\omega x_d/u_c$ values greater than $\pi/2$, the departure in modulus and phase becomes severe for the two cases. The distributed modulus (curve *a*) exhibits resonance phenomena (6) with maximum values at even multiples of $\pi/2$, and minimum values at odd multiples of $\pi/2$. Since $\omega\lambda = 2\pi u_c$, it can be shown that values of $\omega x_d/u_c = \pi/2$ correspond to quarter wave lengths along the ducting; thus $x_d = \lambda/4$. The distributed phase shift (curve *c*) increases by 90-deg intervals between each maximum and minimum of the modulus curve.

Response to a Step-Function Input

The frequency-response plots provided an interesting comparison of lumped and distributed-parameter flow characteristics. In addition to a comparison for a sinusoidal input, a comparison for a step-function input is of interest. The response to a step input provides useful information concerning the transient characteristics, and also will serve to support the results obtained from the frequency-response comparison.

In order to obtain the response to a step input for the lumped-parameter case, it is necessary to utilize Equation [36]. For the same numerical data previously used, Equation [36] becomes

$$\frac{P_2}{A_1}(s) = \frac{9760}{1 + 2.59s} \quad [42]$$

For a unit step function input of A_1, Equation [42] becomes

$$\frac{P_2}{A_1}(s) = \frac{9760}{s(1 + 2.59s)} \quad [43]$$

The inverse transform of Equation [43] is

$$\frac{P_2}{A_1}(t) = 9760\,(1 - \epsilon^{-0.386t}) \quad [44]$$

In order to obtain the response to a step input for the distributed-parameter case, it is necessary to utilize Equation [38]. For the numerical data previously used, Equation [38] becomes

$$\frac{P_2}{A_1}(s) = \frac{9760}{\cosh(0.0282s) + 93.9\sinh(0.0282s)} \quad [45]$$

Letting $0.0282s = \sigma$, and substituting the exponential equivalents for the hyperbolic cosine and sine terms, Equation [45] becomes

$$\frac{P_2}{A_1}(s) = \frac{9760}{\frac{1}{2}(\epsilon^{\sigma} + \epsilon^{-\sigma}) + \frac{93.9}{2}(\epsilon^{\sigma} - \epsilon^{-\sigma})} \quad [46]$$

or

$$\frac{P_2}{A_1}(s) = \frac{205.7}{\epsilon^{\sigma} - 0.97892\epsilon^{-\sigma}} \quad [47]$$

If the denominator of Equation [47] is divided into unity, Equation [47] becomes

$$\frac{P_2}{A_1}(s) = 205.7\sum_{h=0}^{\infty}(0.97892)^h\,\epsilon^{-(2h+1)\sigma} \quad [48]$$

For a step input of A_1, Equation [48] is written as

$$\frac{P_2}{A_1}(s) = \frac{205.7}{s}\sum_{h=0}^{\infty}(0.97892)^h\,\epsilon^{-(2h+1)\sigma} \quad [49]$$

The inverse transform of Equation [49] is

$$\frac{P_2}{A_1}(t) = 205.7\sum_{h=0}^{\infty}(0.97892)^h\,\mu[t - (2h + 1)(0.0282)] \quad [50]$$

where $\mu(t)$ is a unit step function. Thus the time response for the distributed-parameter case is the summation of step-function responses having heights of $205.7(0.97892)^h$, and occurring at varying time delays of $0.0282(2h + 1)$.

For purposes of comparison, values of $P_2(t)/A_1$ versus time for both the lumped and the distributed-parameter case are calculated using Equations [44] and [50]. Response curves are found in Fig. 4. It is seen that the two plots are in fairly good agreement. The distributed response lies slightly below the lumped response; however, both responses reach a steady-state value of 9760.

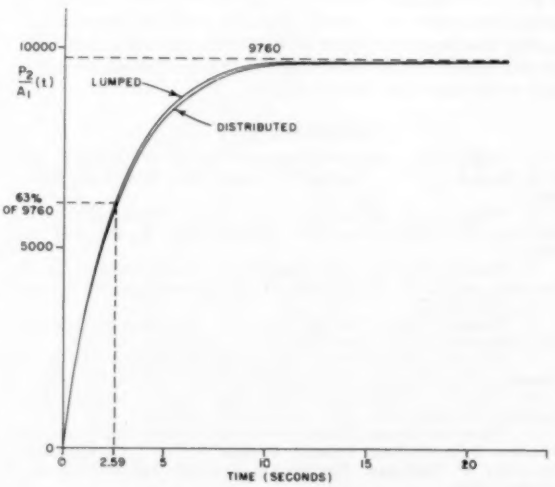


Fig. 4 Time Responses to a Step Function

TABLE 1 COMPARISON OF LUMPED-PARAMETER AND DISTRIBUTED-PARAMETER FLOW CHARACTERISTICS

Time	$\frac{P_2}{A_1}(t)$ lumped	$\frac{P_2}{A_1}(t)$ distributed
0	0	0
1.0	3125	3108
2.6	6170	6095
5.6	8650	8573
11.3	9635	9620
∞	9760	9760

Conclusions

A comparison of lumped-parameter and distributed-parameter flow characteristics has been demonstrated on a frequency-response basis as well as on a step function-time response basis. It was shown that the lumped-parameter representation of ducting air-flow characteristics is valid with certain limitations.

If the ducting section under study was incorporated into a process to be controlled, it is evident from Fig. 3 that control beyond 50 radians per sec would not be possible. Therefore, in the design of the control loop, the crossover frequency should be limited to a value less than 50 radians per sec. If this crossover-frequency limitation is imposed, the lumped-parameter expressions are valid.

It would be desirable to investigate such a control loop by designing the loop in a lumped-parameter basis, with the crossover frequency limited to a value of

$$\omega_c \leq \pi u_c / 9 x_d$$

It is believed that the foregoing criterion for the selection of the limiting value of control loop ω_c is sufficient for most cases. Furthermore, the value of ω_c may be calculated easily without making a frequency-response plot or obtaining the step-function response by inserting the minimum value of u_c and the maximum value of x_d in the foregoing expression for ω_c.

For a case in which the selection of ω_c becomes marginal, then the closed-loop response to a step-function input should be obtained for the lumped-parameter case as well as the distributed-parameter case. This would provide a very complete comparison of the two ducting representations, but is beyond the scope of this paper.

Acknowledgment

The authors gratefully acknowledge the counsel and assistance of Dr. J. T. Trimmer who served as consultant on distributed-parameter analysis work on the Ram Jet Addition, Arnold Engineering Development Center, of which this paper is a part. We are also grateful for the assistance of ARO, Inc., personnel in this work which they have extended (11).

Bibliography

1 "Applied Mechanics—Dynamics," by G. W. Housner and D. E. Hudson, D. Van Nostrand Company, Inc., New York, N. Y., 1950.

2 "Electrical Engineers Handbook," by H. Pender and W. A. Del Mar, John Wiley & Sons, Inc., New York, N. Y., third edition, 1936.

3 "Servomechanisms and Regulating System Design," by H Chestnut and R. W. Mayer, John Wiley & Sons, Inc., New York, N. Y., vol. 1, 1951.

4 "Communication Circuits," by L. A. Ware and H. R. Reed, John Wiley & Sons, Inc., New York, N. Y., 1942.

5 "Electromechanical Transducers and Wave Filters," by W. P. Mason, D. Van Nostrand Company, Inc., New York, N. Y., 1942.

6 "Elements of Acoustical Engineering," by H. F. Olsen, D. Van Nostrand Company, Inc., New York, N. Y., 1947.

7 "Flow and Torque Characteristics of Butterfly Valves at Subcritical and Supercritical Pressure Ratios," by L. C. Garby, et al., University of Michigan, Engineering Research Institute Report, October, 1952.

8 "Flow Characteristics of a Butterfly Valve at Supercritical Pressure Ratios," by L. C. Garby, et al., University of Michigan, Engineering Research Institute Report, June, 1951.

9 "Kent's Mechanical Engineers Handbook—Power Volume," by J. K. Salisbury, John Wiley & Sons, Inc., New York, N. Y., twelfth edition, 1950.

10 "Introduction to the Analysis of Supersonic Ramjet Power Plants," by B. W. Marsh and G. A. Sears, American Rocket Society Paper No. 103-53, December, 1953.

11 "Analysis of Gas Flow Systems for Dynamic Control Purposes," by W. K. McGregor, D. W. Russell, R. W. Messick, and L. F. Burns, Arnold Engineering Development Center, TR-55-11, April, 1956.

Appendix 1

Valve Air-Flow Characteristics

The dynamic relations for the flow of air through a valve are essentially resistive. Experimental research conducted at the University of Michigan (7, 8) has demonstrated that the nozzle equation may be utilized to describe the flow of air through a butterfly valve. The nozzle equation is

$$w = cap_0 \sqrt{\left\{ \frac{2g\gamma}{R\theta_0(\gamma - 1)} \left[\left(\frac{p_1}{p_0} \right)^{\frac{2}{\gamma}} - \left(\frac{p_1}{p_0} \right)^{\frac{\gamma+1}{\gamma}} \right] \right\}} \quad [51]$$

where

c = valve-flow coefficient (0.7 to 1.0)
$\gamma = c_p/c_v = 1.40$
p_1 = downstream static pressure
p_0 = upstream total pressure
θ_0 = upstream total temperature
a = valve open area

For air with $c = 1$, Equation [51] becomes

$$w = \frac{2.055ap_0}{\sqrt{\theta_0}} \sqrt{\left[\left(\frac{p_0}{p_1} \right)^{1.43} - \left(\frac{p_0}{p_1} \right)^{1.71} \right]} \quad [52]$$

The relation between total and static pressure can be written (9)

$$p_t/p_s = (1 - 0.2u^2/u_s^2)^{3.5} \quad [53]$$

where u is the velocity of the air movement, and u_s is the velocity of sound in air. Since the air velocity in the ducting of the plant being considered is very low for all operating conditions, static and total pressures are essentially the same. Therefore, all pressures will be considered to be total pressures.

Utilizing the theory of small perturbations let

$$p_0 = P_0^* + P_0$$
$$p_1 = P_1^* + P_1$$
$$a_1 = A_1^* + A_1$$
$$\theta_1 = \Theta_1^* + \Theta_1$$

Then the total differential of Equation [52] becomes, for $\theta_0 = \theta_1$

$$dw = W(t) = \frac{\partial w}{\partial a_1} da_1 + \frac{\partial w}{\partial p_0} dp_0 + \frac{\partial w}{\partial \theta_1} d\theta_1 + \frac{\partial w}{\partial p_1} dp_1 \quad [54]$$

or in Laplace form

$$W(s) = K_1 A_1(s) + K_2 P_0(s) + K_3 \Theta_1(s) + K_4 P_1(s) \quad [55]$$

where

$$K_1 = \frac{W^*}{A_1^*}$$

$$K_2 = \frac{W^*}{P_0^*} \left[\left(1 - \frac{1.43 \left(\frac{p_1}{p_0} \right)^{1.43} - 1.71 \left(\frac{p_1}{p_0} \right)^{1.71}}{2F^2} \right) \right]$$

$$K_3 = -\frac{W^*}{2\Theta_1^*}$$

$$K_4 = \frac{W^*}{P_1^*} \left[\frac{1.43 \left(\frac{p_1}{p_0} \right)^{1.43} - 1.71 \left(\frac{p_1}{p_0} \right)^{1.71}}{2F^2} \right]$$

$$F = \sqrt{\left[\left(\frac{p_1}{p_0} \right)^{1.43} - \left(\frac{p_1}{p_0} \right)^{1.71} \right]}$$

For values of p_1/p_0 less than 0.528, the function F is constant and equal to 0.257. The designation "choked" or "supercritical" flow is given to operation in this constant F region. When the valve is choked, the downstream pressure no longer has any effect on flow through the valve, and then

$$K_2 = \frac{W^*}{P_0^*} \quad \text{and} \quad K_4 = 0$$

Appendix 2

Ramjet-Engine Air-Flow Characteristics

The flow of air into a ramjet engine may be expressed in terms of the continuity equation, which states that the weight flow is determined by the product of inlet density, area, and velocity (10). Assuming that the engine-inlet nozzle has been removed, stagnation conditions should appear at the inlet—if actual flight conditions are to be simulated.

The engine-flow equation is thus written

$$w_e = \rho a u = \frac{p_2}{R\theta_2}(au) \ldots\ldots\ldots\ldots\ldots [56]$$

Utilizing the theory of small perturbations, let

$$p_2 = P_2^* + P_2$$

$$\theta_2 = \Theta_2^* + \Theta_2$$

Then the total differential of Equation [56] becomes

$$dw_e = W_e(t) = \frac{\partial w_e}{\partial p_2} dp_2 + \frac{\partial w_e}{\partial \theta_2} d\theta_2 + \frac{\partial w_e}{\partial (au)} d(au) \ldots [57]$$

or in the Laplace form

$$W_e(s) = K_3 P_2(s) + K_6 \Theta_2(s) + W_{ed}(s) \ldots\ldots\ldots [58]$$

where

$$K_3 = \frac{W_e^*}{P_2^*}$$

$$K_6 = -\frac{W_e^*}{\Theta_2^*}$$

W_{ed} = engine-flow disturbance.

A Graphical Method for the Analysis of Piecewise Linear Control Systems, With Particular Application to Relay Controls

By R. H. MACMILLAN,[1] CAMBRIDGE, ENGLAND

A graphical method is described for the analysis of piecewise linear control systems. By means of it both transient and steady oscillations can be studied. After the "characteristic curves" of a system have been computed and plotted, it is shown that the effects of dead zone, solid friction, backlash, and various lead devices are easily taken into account, without further computation. The method is illustrated by a more detailed examination of the behavior of relay controls, and a few of the charts obtained by the method are given.

1 Introduction

CONSIDERABLE progress has already been made in the analysis of relay systems; the methods used have been based either on developments of the frequency-response approach or on a solution of the differential equations of motion, originating with the work of Hazen[2] over 20 years ago. In an outstanding paper, with the latter approach, Weiss[3] used graphical constructions in the phase plane to obtain the transient response and, by equating the initial and final states of a cycle, gave charts indicating the conditions for continuous oscillation; he also examined various methods for improving performance. Phase plane techniques have been used by many later writers.

Kahn[4] obtained, by means of the Laplace transformation, a solution for the equations of motion in the form of an infinite series, and used graphical procedures to obtain transient and oscillatory solutions for certain simple systems; the method is ingenious, but is too cumbersome to permit the rapid production of charts useful to a designer. It is, moreover, difficult to obtain an understanding of the physical processes involved. Another approximate solution is due to Bane,[5] but obtaining charts by his method involves much tedious computation.

A fundamental advantage of all differential-equation methods over frequency-response methods is that they readily provide information not only about steady oscillatory states but also about the transient conditions, which are of great importance in many instances.

The method described in this paper was first developed to study relay-control systems. Although this is still the field in which most of the numerical results have been obtained, it soon became clear that the technique can usefully be employed much more widely. The presentation adopted is accordingly a general one, followed by a more detailed application to relay-control systems.

The approach used is to obtain an exact solution of the equations of motion over linear sections; the displacement and velocity (and, where necessary, higher derivatives of the motion) at the beginning of each section are then equated to those at the end of the preceding section. A pair of characteristic curves is thus derived for the system. One relates the velocity at the moment when the displacement is zero to the displacement when the velocity is next zero, i.e., the maximum subsequent displacement; the other characteristic curve relates the initial maximum displacement to the velocity when the displacement is next zero. Each curve thus relates a velocity to a displacement, as does a phase plane trajectory, but the two quantities do not occur simultaneously as in such a trajectory.

By applying simple graphical procedures to these characteristic curves, it is possible to obtain the magnitudes of the successive maxima and the time intervals between them and the zeros during any free oscillation. These points which are precisely determined may be called the "salient points" of the response. The special case in which the conditions at the end of a cycle are the same as they were initially is that of self-sustained oscillation. In this way, one can derive performance charts which are made to be of rather general application by the use of dimensionless variables.

Though not originally so conceived, the method can be regarded as related, in its graphical procedures, to Kahn's approach, but it is capable of giving rather more detailed information about the transient response and is derived quite differently. The characteristic curves of a system are, under certain important conditions, identical in shape to the phase-plane trajectory, and the relationship between them is discussed in the appendix by B. M. Brown.

2 Method

The following sections describe the method in general terms.

2.1 *Systems for Which the Method Can Be Used.* The method can be applied immediately to the analysis of the type of system shown in Fig. 1, provided that the transfer function of the linear elements is such that $Y^{-1}(p)$ is of the first or second degree in p, and the instantaneous relation between the output and input to the nonlinear elements is sectionally linear (as shown, in general, by Fig. 2*h*). This means that at any instant the motion can be described by a linear equation, although the complete motion cannot be so described.

Important special cases directly covered include those shown in Fig. 2, as follows:

---

[1] University Lecturer, Department of Engineering, University of Cambridge. Mem. ASME.

[2] "Theory of Servomechanisms," by H. L. Hazen, *Journal of The Franklin Institute*, vol. 218, no. 3, 1934, pp. 279–331.

[3] "Analysis of Relay Servomechanisms," by H. K. Weiss, *Journal of the Aeronautical Sciences*, vol. 13, no. 7, 1946, pp. 364–376.

[4] "An Analysis of Relay Servomechanisms," by D. A. Kahn, Transactions of the American Institute of Electrical Engineers, vol. 68, part II, 1949, pp. 1079–1088.

[5] "Design Charts for an On-Off Control System," by W. T. Bane, Transactions of the Society of Instrument Technology, vol. 5, no. 2, 1953, pp. 52–71.

Contributed by the Instruments and Regulators Division and presented at a joint session with the Machine Design Division at the Semi-Annual Meeting, Cleveland, Ohio, June 17–21, 1956, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, February 17, 1956. Paper No. 56—SA-17.

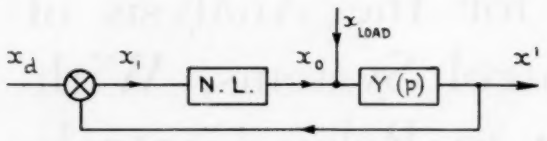


FIG. 1 BLOCK DIAGRAM OF NONLINEAR CONTROL SYSTEM

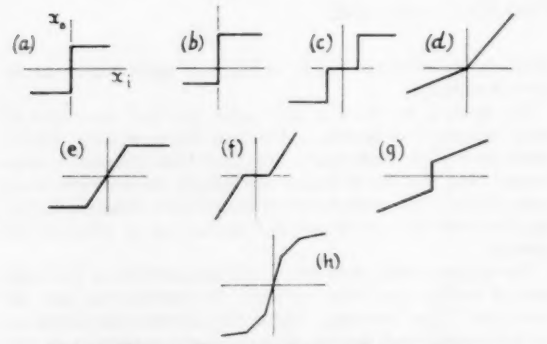


FIG. 2 TYPICAL PIECEWISE LINEAR CHARACTERISTICS

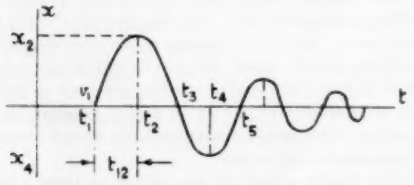


FIG. 3 FORM OF TYPICAL TRANSIENT

(*a*) Relay controls.
(*b*) Relay system with constant output loading.
(*c*) Relay system with dead zone.
(*d*) Linear system with unsymmetrical gain.
(*e*) Linear system with saturation or limiting.
(*f*) Linear system with dead band.
(*g*) Linear system combined with relay action.
(*h*) Approximation to component with variable gain.

It will be shown, further, that the method is applicable to linear or relay systems which involve solid friction, backlash, dead time, or hysteresis, and to systems with velocity saturation or with a combination of two or more of these effects. The method also can be extended to analyze systems in which $Y^{-1}(p)$ is of the third degree in p.

2.2 *General Form of Output Transient.* When the system is disturbed it will subsequently oscillate freely, in general, with a response having the form of Fig. 3, where x_k' is the displacement and v_k' the velocity at any time t_k'. Both the amplitude and period of oscillation vary with time, but the presence of mechanical inertia (or its electrical equivalent) in all practical systems insures that there are no sudden changes of velocity. The slope of the displacement-response shown has, therefore, no discontinuities.

A knowledge of the positions of the successive zeros and maxima defines the shape of the transient sufficiently for all practical purposes. These are the salient points. The method to be described determines the time intervals between successive salient points, the amplitudes of the maxima, and the velocity at each zero.

2.3 *The Characteristic Curves.* For any selected system the forms of the nonlinearity and of $Y(p)$ are known. It is therefore possible to determine, from the equations of motion, the greatest amplitude x_2' achieved when the system begins an oscillation with velocity v_1', as in Fig. 3. To do this it is, of course, necessary to equate the velocity and displacement at each instant where the motion changes from one linear section to another. Similarly, v_3' can be found in terms of x_2', x_4' as a function of v_3', and v_5' as a function of x_4'; x_6' is related to v_5' in exactly the same way as x_2' is related to v_1'.

In general, a solution of the equations of motion over a particular quarter cycle can be expressed as two equations of the form

$$x_2' = f_1(t_{12}') \quad [1]$$

and

$$v_1' = f_2(t_{12}') \quad [2]$$

where $t_{12}' = t_2' - t_1'$. From these relations it is possible to plot the characteristic relating x_2' to v_1' with time interval as a parameter, by selecting suitable values of t_{12}' and calculating the corresponding values of displacement and velocity. This applies to each quarter cycle.

By taking for the units of time and displacement, a time constant T and a length (or velocity V) inherent in the system, it is possible to use dimensionless axes x'/VT and v'/V for the characteristic curves; hereafter the symbol x will be used for x'/VT, v for v'/V, and t for the dimensionless units of time t/T.

2.4 *Use of Characteristic Curves to Obtain Transient Response.* For an instantaneous nonlinearity that is symmetrical about the origin, the quarter-cycle characteristic curves in the third and fourth quadrants must be identical with those in the first and second, so that they might be as in Fig. 4. By equating alternately the velocity or amplitude at the junction of successive quarter cycles, i.e., at the salient points, the complete transient can be determined from any initial conditions. This is easy to do graphically by following the chain line shown in Fig. 4, projecting ordinates to obtain the displacement transient and abscissas to find the velocity transient.

Since there are only two different characteristic curves for a symmetrical nonlinearity, these can be more compactly plotted with both in the first quadrant, as in Fig. 5, where the chain line again determines successive salient points. It will be noted that in Fig. 4 or 5, the time intervals between the salient points are read off directly from the graduations on the characteristics.

2.5 *Characteristics for an Unsymmetrical Nonlinearity.* From Fig. 4 or 5 one can obtain related pairs of values of x, such as x_2 and x_4, with their associated time interval t_{24}, equal to $(t_{23} + t_{34})$. Taking x_{4n} and x_{4n+2} as axes, one can now plot a half-cycle characteristic, relating successive maximum ordinates of the transient, with $t_{4n+2,4n}$ graduated as a parameter along it; a second half-cycle characteristic is obtained similarly to relate such points as x_4 and x_6, with $t_{4n,4n+2}$ as the time parameter; both characteristics are shown in Fig. 6.

To each point (x_{4n}, x_{4n+2}) of the half-cycle characteristic, there corresponds the velocity v_{4n+1} that is attained when the displacement is zero. Its value also can be graduated along the curve together with the time graduations. Half-cycle characteristics for velocity can be plotted similarly to relate v_{4n-1} and v_{4n+1} (with $t_{4n-1,\ 4n+1}$ and x_{4n} graduated along the curve) or to relate v_{4n+1}, v_{4n+3} (with $t_{4n+1,\ 4n+3}$ and x_{4n+2} as parameters).

The chain line in Fig. 6 indicates how the successive maxima of a transient can be determined rapidly from a half-cycle displacement characteristic, for a system with an unsymmetrical nonlinearity.

It is also possible to plot full-cycle characteristics and double-cycle or multiple-cycle characteristics. Half-cycle characteristics

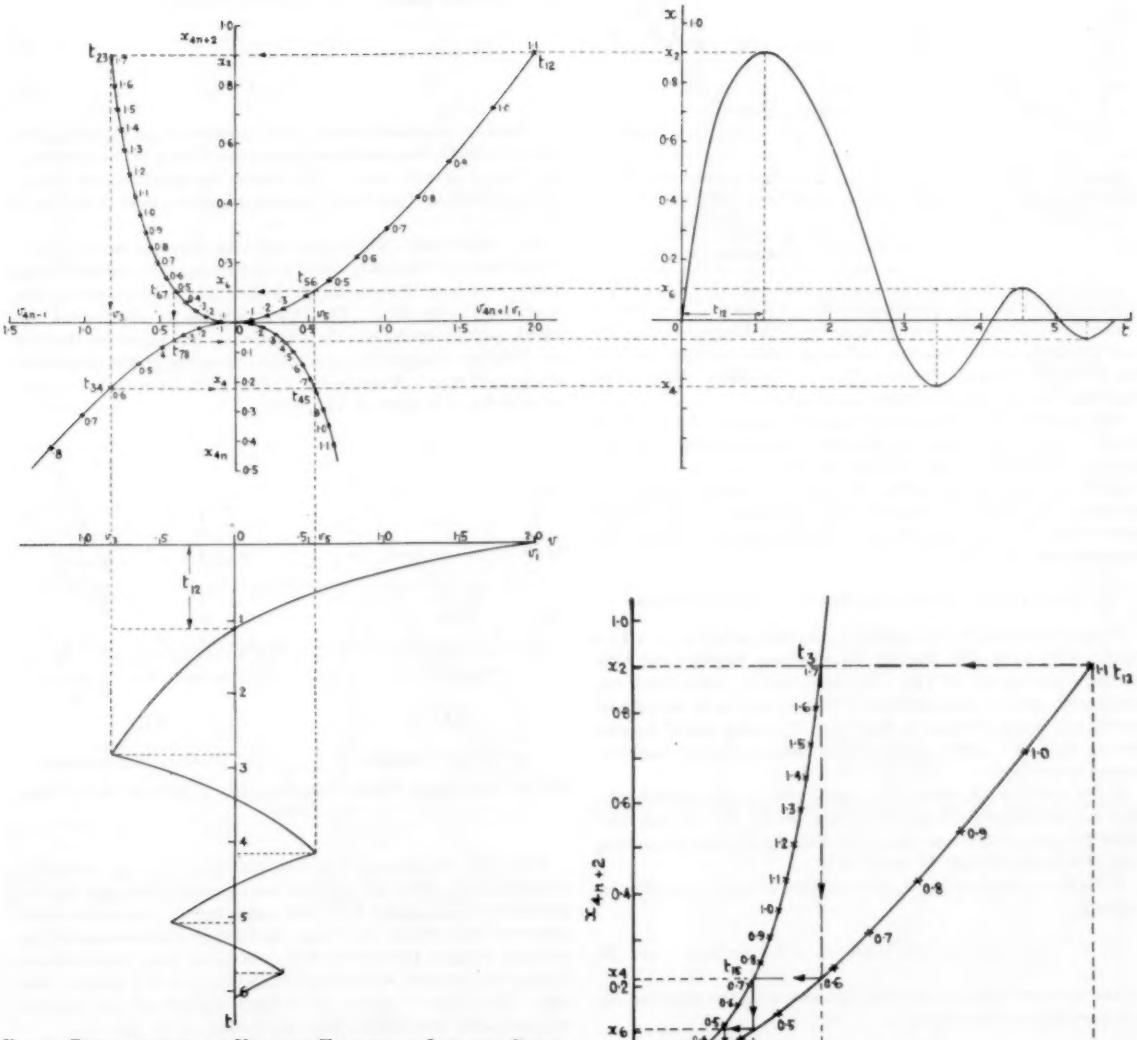


FIG. 4 DISPLACEMENT AND VELOCITY TRANSIENTS OBTAINED FROM QUARTER-CYCLE CHARACTERISTICS

FIG. 5 IMPROVED METHOD OF OBTAINING TRANSIENT FROM QUARTER-CYCLE CHARACTERISTICS

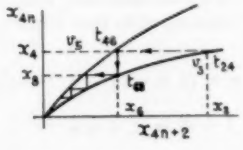


FIG. 6 OBTAINING A TRANSIENT FROM HALF-CYCLE CHARACTERISTICS

are particularly useful for studying systems with a constant-velocity input or those with solid friction. For a symmetrical nonlinearity there is only one half-cycle characteristic and for an unsymmetrical one there is only one full-cycle characteristic. Multiple-cycle characteristics are needed to investigate subharmonic phenomena.

2.6 *Sustained Oscillation.* It is quite possible for the curves of Fig. 5 to cross each other, as in Fig. 7 or in Fig. 8. In a system with dead time, for example, the velocity at the end of a half cycle may exceed that at the beginning of it. In such a case, the chain line defining the transient converges upon the point S, with amplitude x_s, irrespective of whether the initial amplitude was greater or less than x_s. When the characteristics cross more than once, the intersections represent alternately stable and unstable states of sustained oscillation. In Fig. 8, for example, the system comes to rest at $v = 0$ after a disturbance of amplitude less than x_u, but builds up an oscillation of amplitude x_s following any disturbance greater than x_u.

2.7 *Application to Higher-Order Systems.* If $Y^{-1}(p)$ is of the third degree in p, it is necessary to equate values of acceleration (in addition to those of displacement and velocity) at the salient

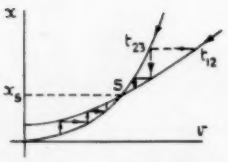


FIG. 7 QUARTER-CYCLE CHARACTERISTICS OF OSCILLATORY SYSTEM

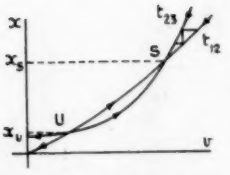


FIG. 8. STABLE AND UNSTABLE OSCILLATORY STATES
(S is stable point; U is unstable point.)

points, which greatly complicates the problem. To obtain a transient solution, two sets of quarter-cycle characteristics are now required, one for constant values of initial acceleration and the other for constant values of final acceleration. With these available the method of solution is obvious.

The solution for sustained oscillation is simpler, the key to it being to plot half-cycle or full-cycle characteristics with the property that the initial and final accelerations are equal (but opposite in sign for the half-cycle case). In certain special instances it is possible to calculate the acceleration at the salient points from the physical nature of the system, and this greatly simplifies the analysis.

3 APPLICATION TO RELAY-OPERATED CONTROL SYSTEMS

We shall now apply the method to a system as in Fig. 1, with a relay switch as in Fig. 2(*a*) for its nonlinear element and with $Y^{-1}(p)$ equal to $p(1 + Tp)$. This is a control system with one integration and one time constant T (corresponding to the ratio of inertia to viscous friction) in the loop. The relay switch insures always that the motor applies maximum positive or negative torque to the load.

If the relay switch were to be locked over in one position, the load would accelerate until the motor torque M, and the frictional torques $\mu\dot{x}'$ became equal, when the load would be moving with maximum velocity V, equal to M/μ.

With the symbols of Fig. 1, the equation of motion can thus be written

$$p(1 + Tp)x' = V \text{ sign } (x_d - x') - x_{\text{load}} \quad [3]$$

In the absence of input x_d or disturbance x_{load} the equation for free oscillation is therefore

$$p(1 + pT)x' = \pm V \quad [4]$$

the sign being negative when x' is positive, and vice versa, as a result of the switching action. This system is the same as that studied by previous writers and comparison of results is thus possible. It is the simplest approximation to practical systems from which significant results are obtainable.

3.1 *The Quarter-Cycle Characteristics.* If x_1, v_1 and x_2, v_2 are the values of x and v at times t_1 and t_2, then solution of Equation [4] for positive x gives

$$x_2 - x_1 = (1 + v_1)(1 - e^{-t_{12}}) - t_{12} \quad [5]$$

and

$$1 + v_2 = (1 + v_1)e^{-t_{12}} \quad [6]$$

For the first quarter cycle, $x_1 = 0$ and $v_2 = 0$, so

$$v_1 = e^{t_{12}} - 1 \quad [7]$$

$$x_2 = e^{t_{12}} - 1 - t_{12} \quad [8]$$

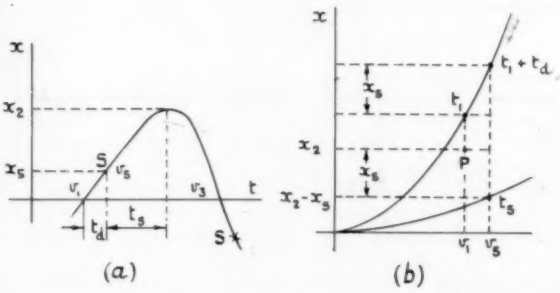
For the second quarter cycle, we obtain similarly

$$-v_2 = 1 - e^{-t_{23}} \quad [9]$$

$$x_2 = e^{-t_{23}} - 1 + t_{23} \quad [10]$$

As would be expected from physical considerations, Equations [9] and [10] can be obtained from Equations [7] and [8] by substituting $-t_{23}$ for t_{12} and v_2 for v_1. The form of the quarter-cycle characteristics obtained from these equations is shown to scale in Figs. 4 and 5.

3.2 *Dead Time.* The presence of dead time t_d in the relay (i.e., a finite time t_dT elapses between zero displacement and switching) causes switching to occur during the first and third quarter cycles, as shown in Fig. 9(*a*). The characteristics for the second and fourth quarter cycles are thus unaltered, but those for the first and third are changed by an amount depending on the magnitude of the dead time. A method of finding them from the characteristics of Fig. 5 is given in Fig. 9(*b*).



(*a*) Form of transient (*b*) Graphical construction

FIG. 9 OBTAINING CHARACTERISTICS FOR A SYSTEM WITH DEAD TIME

When the displacement is zero, the velocity is v_1; switching occurs t_d later, when the velocity has increased to v_s and the displacement is x_3, found from the upper curve. From the lower curve we now obtain $(x_2 - x_s)$, the further displacement till the velocity is again reduced to zero, and t_s the time which elapses during this process; the actual displacement x_2 is x_s greater than this. The point P at (v_1, x_2) is thus a point on the required characteristic and has the time graduation $(t_s + t_d)$.

The characteristics for various values of t_d are shown to scale in Fig. 10. By using the two curves appropriate to any value of t_d, the salient points of a transient can be found, and the amplitude x_a and period $t_p = 2(t_{12} + t_{23})$ of continuous oscillation are shown clearly by their intersection.

3.3 *Backlash and Dead Zone.* If the relay has the input-output relation of Fig. 11, it suffers from backlash of dimensionless magnitude x_b. The transient for a system with backlash can be found by shifting the quarter-cycle characteristics relatively to each other, in the x-direction, by an amount x_b, as in Fig. 12. Since the curves now intersect, the system oscillates continuously with amplitude x_a, as it does in the presence of dead time.

If there is a spacing x_eVT between the relay contacts, with no restoring force applied to the system in this dead zone, the relay action is that shown in Fig. 2(*c*). Since, with the system we are considering, the loss of velocity in crossing the dead zone is always the same, irrespective of the entry velocity, the presence of dead zone can be taken into account by shifting the quarter-cycle curves relative to each other by an amount x_e in the v-direction, as in Fig. 13. In calculating the periods and settling times

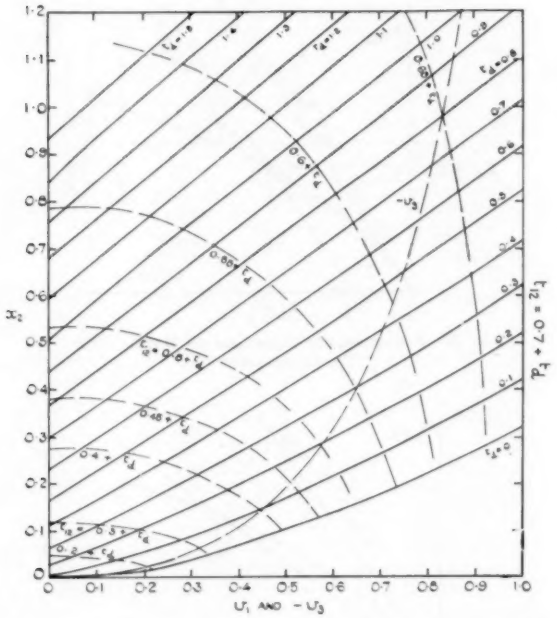


FIG. 10 QUARTER-CYCLE CHARACTERISTICS FOR SYSTEM WITH DEAD TIME

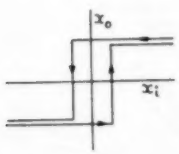


FIG. 11 CHARACTERISTIC OF COMPONENT WITH BACKLASH

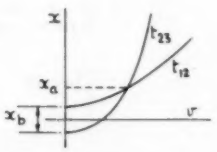


FIG. 12 OBTAINING TRANSIENT OF SYSTEM WITH BACKLASH

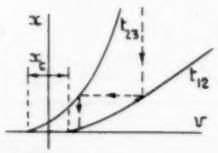


FIG. 13 OBTAINING TRANSIENT OF SYSTEM WITH DEAD ZONE

for a system with dead zone, allowance must be made for the time taken to cross the dead zone (which is not independent of entry velocity). No matter how great the initial disturbance may be, a system with pure dead zone comes to rest after a finite number of overshoots.

It is possible to obtain transients and amplitudes of sustained oscillation for systems with both backlash and dead zone or dead time, by taking the appropriate curves and shifting them relatively to each other in either or both directions. A convenient procedure is to trace one of the characteristics on an overlay which can be placed above the other one and moved over it. In this way charts have been obtained showing, for instance, the amount of dead zone needed to insure that not more than one overshoot shall occur in a system with a given amount of backlash.

Two typical charts obtained by this method are reproduced: Fig. 14 indicates the number of overshoots and time to settle for a system with any amount of dead zone, and Fig. 15 shows the amplitude and period of oscillation of systems with any combination of dead time and backlash. Both charts are drawn for dimensionless variables.

3.4 *Half-Cycle Characteristics.* The derivation of half-cycle characteristics has been mentioned in Section 2.5. Displacement characteristics obtained as there indicated are shown to scale in Fig. 16 for various dead times; as the nonlinearity is symmetrical, there is one curve only for each value of dead time $t_d T$. The maximum amplitudes of a transient in such a system can be obtained directly from one of these curves by using the construction of Fig. 17; lines are projected parallel to each axis alter-

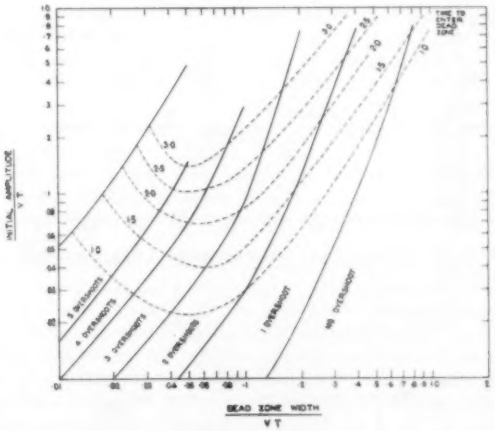


FIG. 14 OVERSHOOTS AND SETTLING TIME OF RELAY SYSTEMS WITH DEAD ZONE

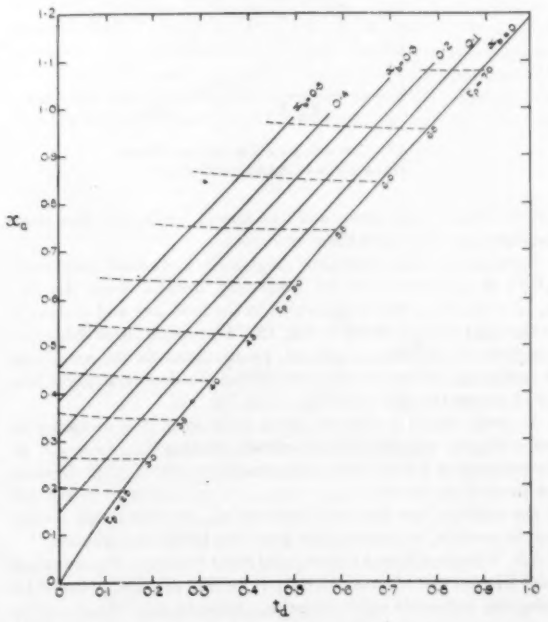


FIG. 15 AMPLITUDE AND PERIOD OF OSCILLATION OF RELAY SYSTEMS WITH DEAD TIME AND BACKLASH

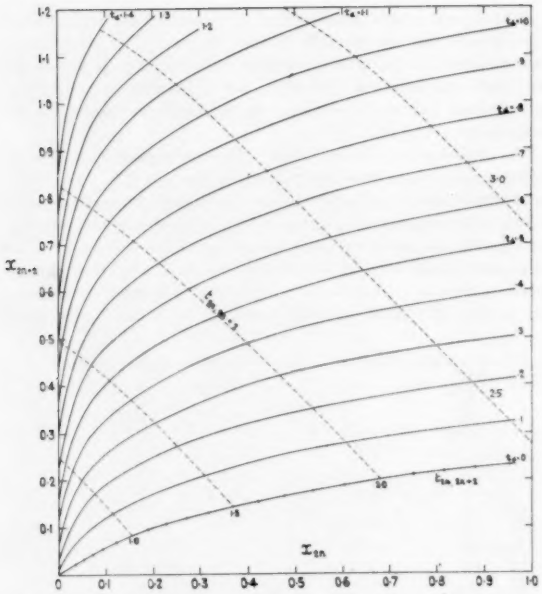


FIG. 16 HALF-CYCLE DISPLACEMENT CHARACTERISTICS

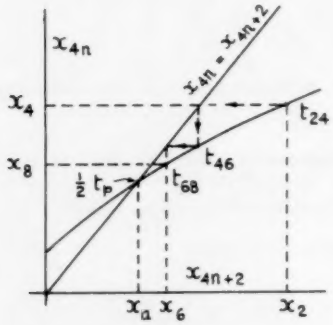


FIG. 17 OBTAINING A TRANSIENT FROM HALF-CYCLE CHARACTERISTICS

nately between the curve and the line $x_{4n} = x_{4n+2}$. Sustained oscillation occurs where these two meet.

Backlash of dimensionless magnitude x_b (actual magnitude x_bVT) is accounted for by projecting instead from the line $x_{4n+2} = x_{4n} + x_b$, which is parallel to the first line and distant x_b to the right of it as shown in Fig. 18. Hysteresis (here defined as backlash of variable magnitude, proportional to the amplitude of oscillation) is accounted for by tilting the sloping straight line, which passes through the origin, as in Fig. 19.

To treat similarly a system with dead zone, it is necessary to use half-cycle velocity characteristics, relating v_{2n+1} to v_{2n-1}. In the presence of a dead zone of dimensionless width x_c, projections are made from the line $v_{4n+1} = v_{4n-1} - x_c$, as shown in Fig. 20. If the width of the dead zone exceeds x_{cm}, no continuous oscillation is possible, no matter how great the initial disturbance.

3.5 *Constant Output Loading and Solid Friction.* The effects of solid friction and of constant output loading are readily found by using the half-cycle characteristics. Solid friction effectively increases the restoring force during the first and third quarter cycles, while decreasing it on the second and fourth (see Fig. 21).

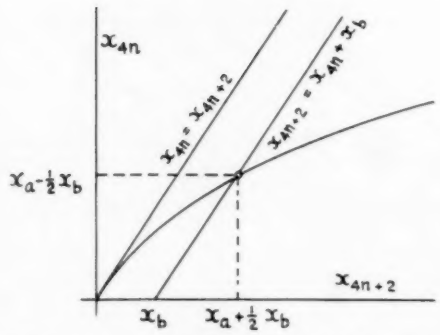


FIG. 18 ALLOWING FOR BACKLASH ON HALF-CYCLE CHARACTERISTIC

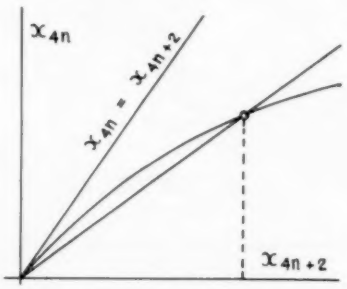


FIG. 19 ALLOWING FOR HYSTERESIS ON HALF-CYCLE CHARACTERISTIC

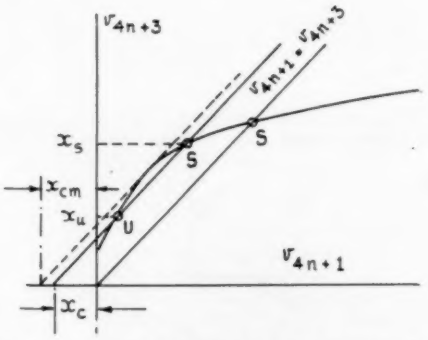


FIG. 20 ALLOWING FOR DEAD ZONE ON HALF-CYCLE CHARACTERISTIC

If the amount of the friction is defined by r_f, the ratio of the frictional force to the restoring force, then the net restoring force is either increased in the ratio $(1 + r_f)$ or decreased in the ratio $(1 - r_f)$. Thus the characteristics for a frictionless system can be used, provided that the scales of x'/VT and v'/V are altered in the foregoing ratios to allow for the effective change in V.

From suitably scaled quarter-cycle curves, a series of half-cycle displacement characteristics is thus obtained and shown to scale in Fig. 22, for various values of r_f. A transient for a system with both friction, defined by $r_f = 0.2$, and backlash $x_b = 0.3$ is found by constructing the chain line shown.

The result of output loading (Fig. 2*b*) is rather similar, for it effectively increases the restoring force in the ratio $(1 + r)$ when the displacement is positive, and reduces it in the ratio $(1 - r)$ on the second half cycle. For each value of r, a pair of half-cycle

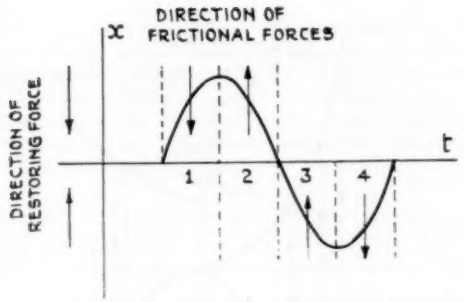


FIG. 21 EFFECT OF SOLID FRICTION ON NET RESTORING FORCE

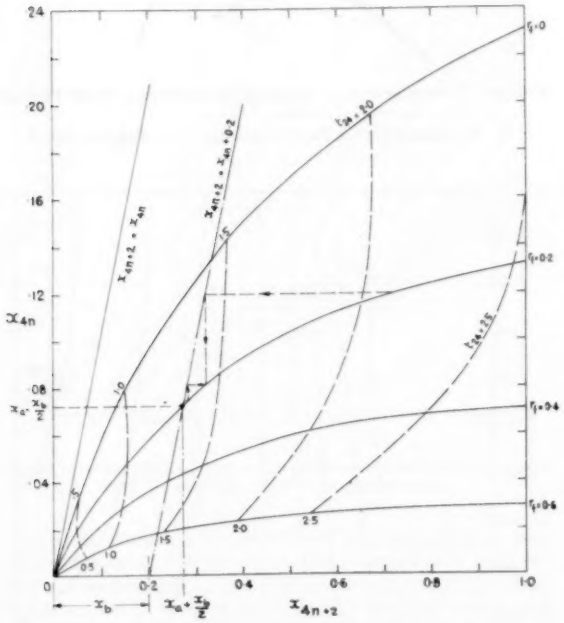


FIG. 22 HALF-CYCLE CHARACTERISTICS FOR SYSTEMS WITH SOLID FRICTION

velocity curves can therefore be obtained by scaling directly from the no-load half-cycle velocity characteristic. Alternatively, half-cycle displacement characteristics can be obtained, in the manner of finding those for solid friction, from suitably scaled quarter-cycle curves. From such characteristics one can obtain the amplitude of continuous oscillation and the "droop" or steady-state error caused by the loading. As shown in Fig. 23(*a*) this is half the difference of the two half-cycle amplitudes during continuous oscillation, and it can thus be measured directly from the figure.

The effect of a constant input velocity is mathematically identical with loading the system.

3.6 *Lead Devices, and Others, to Improve Performance.* We shall now consider briefly the use of this method to analyze the effect of certain devices used in practice to improve the performance of relay controls. It will be found that the technique helps to give an insight into the operation and relative effectiveness of these devices.

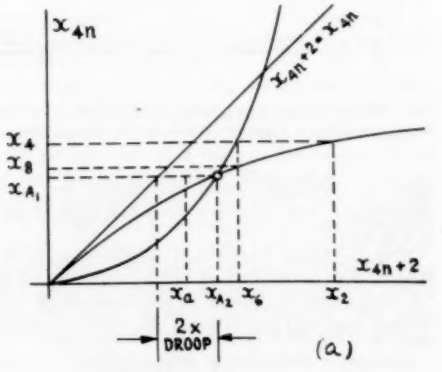
A "dead-zone brake" increases the effective width of the dead zone, without increasing the permanent error by enlarging the actual width. The result is to reduce the amplitude of oscillation or eliminate it entirely. Analysis by the usual method for dead zone can be used.

"Flip-flop" or "front-lash" units effectively introduce negative backlash. Analysis by the usual method for backlash (Fig. 19) will show the amount of front-lash needed to eliminate continuous oscillation.

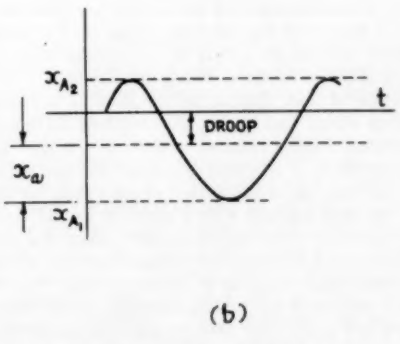
"Variable lead devices" insure that switching occurs before the displacement is reduced to zero. The relay operates when there is a lead given by $x = f(-v)$. The effect of this lead can be studied directly on the quarter-cycle characteristics, for arbitrary $f(-v)$, as shown in Fig. 24. The lead curve is plotted on an overlay and, instead of projecting parallel to the v-axis to obtain the transient, the intersection of the lead curve with the quarter-cycle characteristic is the point used.

It is clear from a consideration of the third quadrant of Fig. 24(*a*) that the system will have $x_4 = 0$ (i.e., no overshoot) if $f(-v)$ has exactly the form of the third (or first) quarter-cycle characteristic, which is nearly parabolic. Parabolic lead has been used in practice, but it is difficult to apply it without undue complication; simpler to use are either of the straight-line approximations to the ideal lead characteristic, as shown in Fig. 25. The first of these (line *a*) provides a lead proportional to velocity; that is

$$x + t_l v = 0 \qquad [11]$$

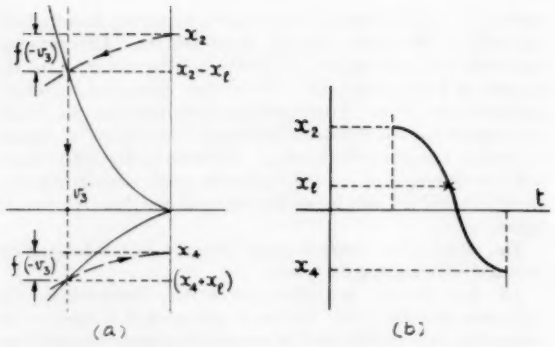


(a) Half-cycle characteristics



(b) Form of oscillation

FIG. 23 OBTAINING AMPLITUDE OF OSCILLATION AND STEADY ERROR OF LOADED SYSTEM



(*a*) Construction of quarter-cycle characteristics (*b*) Form of transient

FIG. 24 OBTAINING THE TRANSIENT WITH ARBITRARY LEAD

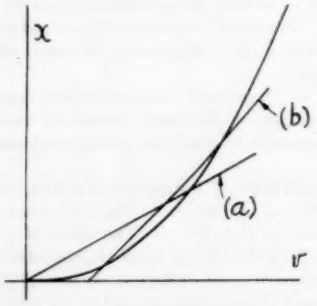


FIG. 25 LINEAR APPROXIMATIONS TO "IDEAL" LEAD CHARACTERISTIC

where t_l is the lead-time constant; this type of lead is considered in more detail in the next paragraph. A better approximation to the ideal characteristic is a relation of the form

$$x + t_l v + k(\text{sign } v) = 0 \quad \ldots\ldots\ldots\ldots [12]$$

as shown by line (*b*) in Fig. 25. Such a lead function was proposed by Weiss and it is clear from our method of analysis why it is advantageous. When using it, one can obtain zero overshoot for a relatively large disturbance, together with a better speed of response than is given by the simple proportional lead.

"Proportional lead," with the moment of switching defined by Equation [11], is the simplest lead device which is effective; the characteristic is line (*a*) of Fig. 25 with a slope t_l. For a system with dead time, the salient points of a transient with initial amplitude x_2 is obtained as shown in Fig. 26, line (*a*).

As the amount of lead, defined by the slope t_l, is increased, switching occurs earlier and earlier, the overshoot from a given amplitude eventually being reduced to zero. If t_l is still further increased, there will be a tendency to undershoot but switching will again occur before the displacement changes sign, the system returning to the zero position with a series of torque reversals and without overshoot. For each value of t_l, there is a limiting amplitude below which there is no overshooting; this is roughly analogous to the critical damping of linear systems. The values of this critical amplitude for various amounts of lead and dead time are shown in Fig. 27. For this system it can be shown that provided

$$t_l > t_d + 0.3 \quad \ldots\ldots\ldots\ldots [13]$$

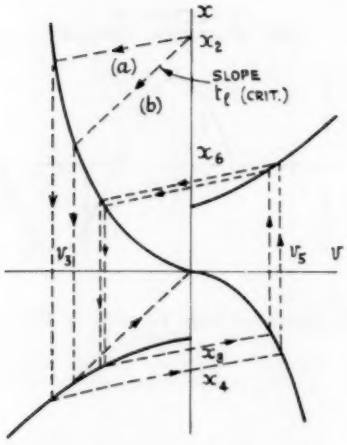


FIG. 26 CONSTRUCTION OF TRANSIENTS IN SYSTEM WITH PROPORTIONAL LEAD

(*a*, A transient for oscillatory system; *b*, critical transient.)

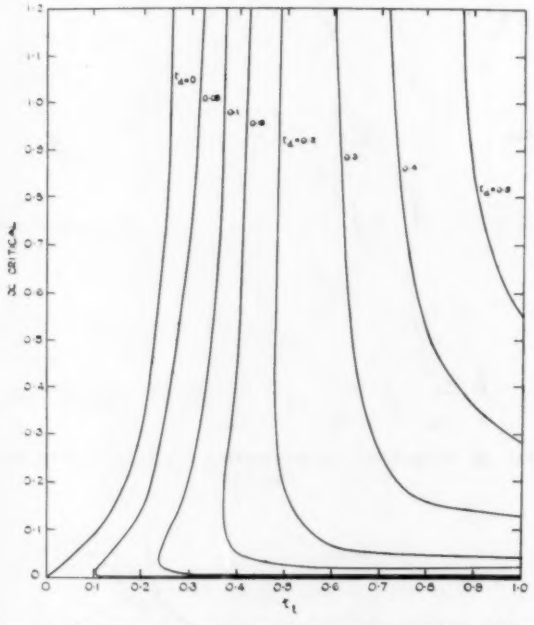


FIG. 27 CRITICAL AMPLITUDE OF DISTURBANCE TO AVOID OVERSHOOT WITH VARIOUS AMOUNTS OF LEAD AND DEAD TIME

then no overshoot occurs, no matter how great the initial disturbance; this relation defines the asymptotes of the curves in Fig. 27.

If the transient is such that eventually $x_{4n+2} = x_{4n}$ in Fig. 26, then continuous oscillation will be sustained at that amplitude. The amplitude and period of such oscillations for various values of dead time and proportional lead are plotted in Fig. 28.

4 CONCLUSION

The method described has now been developed sufficiently to provide charts which can be used to assist the designer, and ex-

amples of these have been given. The basic simplicity of the method is thought to lead to a useful insight into the physical processes involved, a thing which it is difficult if not impossible to obtain from more elaborate mathematical analysis or from analog techniques, but these should both undoubtedly be used in conjunction with it. No approximation beyond that inherent in any graphical procedure is involved in the method.

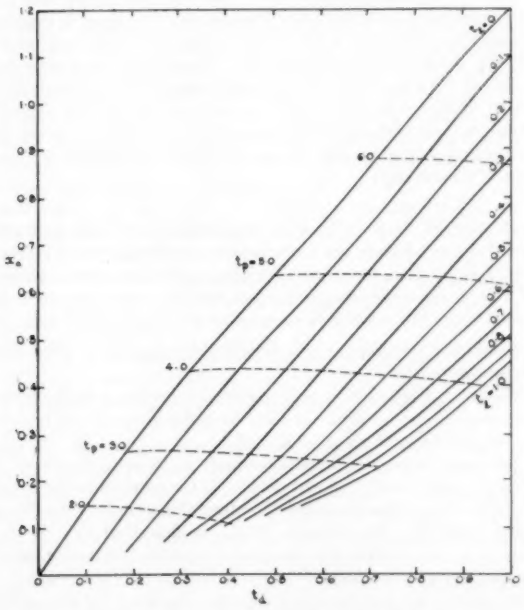


FIG. 28 AMPLITUDE AND PERIOD OF OSCILLATION OF RELAY SYSTEMS WITH DEAD TIME AND PROPORTIONAL LEAD

Once the characteristic curves have been computed for any particular system, various graphical techniques have been devised to take into account such effects as the presence of backlash, dead zone, and so on; it is relatively easy to devise further procedures to help in considering such factors as an exponential time delay for the build-up of torque after switching, or the effect of output loading. To solve the latter problem, for instance, plotting of the half-cycle characteristics to logarithmic scales greatly facilitates changing the scale, since this can then be achieved by simply shifting the axes, a technique which there has not been space to describe in detail.

It is hoped, however, the description given has been sufficient to indicate that the complexity of practical systems is within the scope of this method. In particular, it could readily be applied to the analysis of process controls if the transfer function of the process is approximated, as suggested by Hrones and others,[6] by a combination of an exponential delay and dead time. It might also be applied to the study of sampling systems and those with step-by-step control. If the quarter-cycle and half-cycle characteristics were computed for a selection of standard systems (the number of possibilities is limited), designers could then obtain rapidly from them the particular charts they required. The most pressing need now is to compute characteristics for linear systems with various types and combinations of saturation or limiting.

[6] "On the Automatic Control of Generalized Passive Systems," by Kun Li Chien, J. A. Hrones, and J. B. Reswick, Trans. ASME, vol. 74, 1952, pp. 175–185.

Appendix

Relation Between Characteristic Curves and Phase Plane

BY B. M. BROWN,[7] GREENWICH, ENGLAND

If we draw a phase diagram for a motion of the type represented by Fig. 3, the salient points correspond to the points where a particular trajectory cuts the x and v-axes. Also, the distances of these points from the origin are, respectively, v_1, x_2, v_3, . . . (see Fig. 29). If we consider one of the characteristic curves of Fig. 4, say, the one in the first quadrant, this is seen to be the locus of points (v_1, x_2), (v_3, x_4) . . .; in other words points whose coordinates are the intercepts on the axes of the positions of the trajectories in the first quadrant of the phase plane. In order to plot this characteristic curve from the phase-plane diagram it is only necessary to read off these intercepts from the latter, one point being obtained from each trajectory drawn. The same is true for the other three characteristic curves. (Notice, however, that the characteristic curve in the *second* quadrant is obtained from the trajectories in the fourth quadrant of the phase plane and vice versa.)

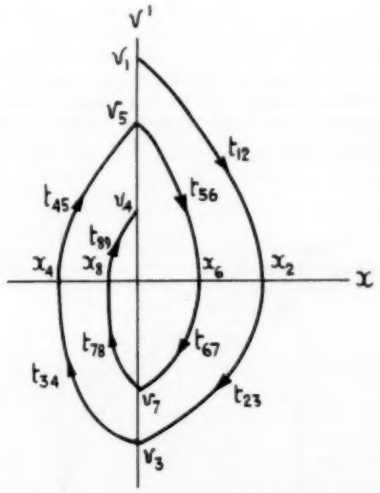


FIG. 29 TRAJECTORY ON PHASE PLANE

The relation between these two graphical representations is particularly interesting in the important special case when the differential equations in the two regimes have the form $F(\dot{v}, v) = 0$, x not appearing explicitly. In this case the general solution will take the form $\phi(v, x + C) = 0$. The phase plane trajectories are therefore families of curves identical in shape and size and differing only, one curve from another, by a displacement parallel to the x-axis. If the trajectories are all displaced in this way so that they pass through the origin, each family reduces to a single curve as in Fig. 30. This figure shows how the characteristic values v_1, x_2, v_3, x_4 . . . of the motion are related, and it is clear if the diagram is rotated through a right angle in a clockwise sense that Fig. 4 is obtained. Thus, in the special case we are now considering, the characteristic curves are identical in shape and size with the phase plane trajectories.

[7] Royal Naval College.

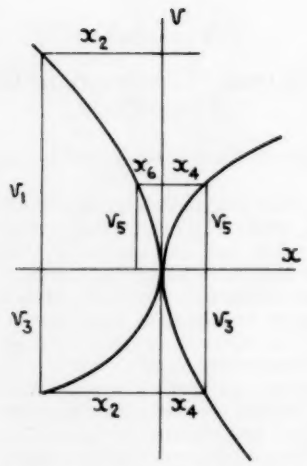


FIG. 30 PHASE PLANE TRAJECTORIES DISPLACED TO PASS THROUGH THE ORIGIN

This is not true if the differential equations contain x explicitly. When the trajectories are displaced to pass through the origin in this case, they no longer coincide to form a single curve but give instead two infinite families that touch at the origin. The loci corresponding to the characteristic curves are obtained by taking one point from each curve in each quadrant, the coordinates of this point being the intercepts on the axes of the corresponding original trajectory. Fig. 31 shows how the

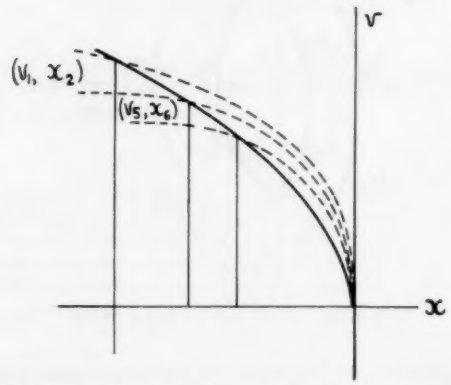


FIG. 31 CHARACTERISTIC CURVE RELATED TO PHASE PLANE TRAJECTORIES

characteristic curve corresponding to the first quadrant trajectories is obtained.

It is clear from the foregoing that there is a close connection between the characteristic curves and the phase plane. But, whereas the complete phase plane will provide all the information that can be deduced from the characteristic curves, the latter, being single curves rather than families, can be drawn much more easily. Moreover they provide their information in more compact form and allow such phenomena as limit cycles to be identified easily. However, the phase plane is a most useful tool for deriving the properties of the characteristic curves, particularly when taking account of dead time, backlash, and the like.

Discussion

J. L. SHEARER.[8] The graphical analysis described in this paper gives a good basic understanding of the performance of simple processes under the action of on-off control and it gives the reader a particularly satisfying feeling about how the fundamentals of transient response can be applied to systems of this kind. However, it would appear that the author's graphical solution would be very difficult to achieve with processes higher than second order because of the extra complications involved in evaluating and keeping track of initial conditions when shifting from one type of linear response to another. The results the author has collected from a number of analyses of simple systems seem to be worth while not only to the young engineer who is just learning about control systems but also to old hands who can relate the limitations found in very simple analyses to systems of much greater complexity.

ROBERT W. BASS.[9] This paper provides a welcome and very useful contribution to the subject. Because the graphical method described by the author provides information about the transient as well as the steady-state response, and for other reasons mentioned later, the method is superior to the approximate Describing-Function technique which has also been applied to piecewise-linear systems (1).[10]

However, the writer is perhaps not presenting a disinterested appraisal of the method—in fact, he himself developed it independently nearly three years ago (2).

Because this graphical method simply picks out a succession of what the author calls "salient points" in the corresponding phase-space, the writer had proposed (2) calling this technique the phase-graph method. In October, 1953, he described (2) "a new method of calculating the transient response, hereinafter called the phase-graph method. [This] method permits the reading of the duration, amplitude, and maximum velocity of the nth oscillation of a [relay] servosystem. . . ." The writer then noted (2) that Kahn's method (3) required a distinct graph to be plotted for each distinct dead-zone width, but that in his own phase-graph method "a single graph suffices for the entire range of dead-zones. . . . [The] graph is not [yet] ready, but will appear in a later report."

In November, 1953, the writer left this project for about a year. Before doing so he left with Dr. J. M. Kopper (the Project Director) and Dr. H. S. McDonald written descriptions of how to plot and use graphs identical with those of Figs. 7, 8, 9, 12, and 13 of the author. In March, 1954 (4) and in subsequent reports to Frankford Arsenal, these graphs were presented by Dr. McDonald who developed a more suggestive physical terminology for them: the upper curve of Fig. 4 was called the "acceleration curve" and the lower, the "deceleration curve." Dr. McDonald also noticed the connection with the standard phase-plane plot which is pointed out by B. M. Brown in the Appendix. This led Dr. McDonald to observe (5) that the method could easily be extended to nonlinear motors whose transient response is not described by Equation [4] but for which the acceleration and deceleration curves can be found empirically. In reports during 1954 Dr. McDonald also presented graphs giving the same information (for example) as that obtained in the author's Figs. 15 and 28. The writer and Dr. McDonald were also familiar with Equation [13], and the writer had found an analytical expression of which the information presented in Fig. 27 is a special case.

---

[8] Assistant Professor of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Mass. Mem. ASME.

[9] Staff Scientist, RIAS, Inc., Baltimore, Md.

[10] Numbers in parentheses refer to the Bibliography at the end of this discussion.

In the Final Report of this project (6), the writer devoted a chapter to his Phase-Graph Method. In illustration, it was applied to the author's Equation [4], where the relay was assumed to possess simultaneously dead-zone, "hysteresis" (i.e. the author's "backlash"), and time-delay. (Thus the relay characteristic was given by combining the author's Fig. 2(*a*) or (*b*) with his Fig. 11.) Results identical with those of the author were obtained.

The author does not explicitly mention it, but the phase-graph method is of considerable help in analytical investigations. For example, by eliminating t_{12} from [7] and [8], and by applying some analytic geometry to Figs. 7, 8, 12, and 13, the writer obtained expressions which give an exhaustive description of the steady-state behavior of the relay servo for all ranges of the four parameters: dead-zone width, hysteresis, time-delay, and the coefficient of linear or quadratic error-rate feedback. (This last coefficient corresponds to the author's t_1 of Equation [11].) These results have been summarized in several places (1, 7).

Incidentally, references (1, 6, 7) present a new method of stabilizing relay servos, based on reference (8), which is superior to the conventional "variable-lead" devices discussed by the author in section 3.6. For example, stabilization by Equation [11] or [12] usually introduces "chattering," which the writer's method (1, 7) does not.

In 1955, D. F. Lawden published (in his book "Mathematics of Engineering Systems," Methuen, London; cf. pp. 295, 297) some special cases of the phase-graph idea, but did not discuss the general possibilities of the method.

Also in 1955, the writer stated (in Section 5.1, *Piecewise Linearity and the Iteration of a Transformation*) (6) that "the phase-graph method can prove useful in almost any problem of *piecewise-linear mechanics*. For by definition the phase plane (space) can be broken up by lines (surfaces) into regions in which the differential equation, nonliner *in the large*, is linear. Now if the coefficients are constant, the linear differential equations can be integrated easily to give the phase-trajectories. Each family of trajectories can be considered as a transformation T which transforms points $P = (e, \dot{e})$ from one boundary line (surface) into the corresponding points $T(P)$ on the subsequent boundary. Now by applying these transformations successively, or (in the cases of symmetry) by iterating them, a sequence of points on the phase path will be obtained which gives a general description of the shape of the trajectory.

"If the successive points on any particular line or surface converge to a limit point P_0, this point P_0 will have the property that $T(P_0) = P_0$. In other words, P_0 is a *fixed point* of the transformation T; by the uniqueness of trajectories in phase space, such a point will provide the initial conditions of a closed integral-curve. This method for finding limit-cycles . . . is classical in celestial mechanics."

This description applies, e.g., to the transformation defined by Fig. 17, where x_a is the fixed-point giving a limit-cycle.

A very interesting discussion of the possible nature of such transformations T and their fixed-points is given by R. E. Kalman in a paper to appear in the Proceedings of the Second Symposium on Nonlinear Circuit Analysis, Brooklyn Polytechnic Institute.

In the report mentioned (Ch. VIII, Fourier Series, Bifurcation Equations, and the Methods of Equivalent Linearization) (6) the writer compared phase-graphs with the popular describing-function method: "Our phase-graph method gives the entire transient and steady-state response, furnishing at a glance the number, size, and duration of overshoots, and the location and stability of limit-cycles. Moreover, the phase-graph method is theoretically exact and requires the plotting of only two loci. Since the frequency-response approximation also requires the plotting of two loci and since it establishes no more than the location and stability of limit-cycles, it would seem that it can compete with the phase-graph method solely because it is easier to derive and to use for synthesis. But in fact, the phase-graph method can be extended only to other piecewise-linear devices, while the frequency-response method . . . is subject to no such restriction. On the other hand, frequency response can be used only for synthesis of (usually linear) compensating—not optimizing—networks. The phase-graph method can be used for the synthesis of linear or nonlinear devices, both for compensation and optimization. . . ."

Since then the writer has further developed the phase-graph method. For example, he has proved that the phase-portrait of a wide class of continuous nonlinear differential equations is topologically equivalent to that of a "sufficiently close" piecewise-linear approximating equation; hence the phase-graph technique can, in principle, be applied to almost any nonlinearity. These and other results will appear in his book Nonlinear and On-Off Control Systems, to be published by Princeton University Press in 1957.

BIBLIOGRAPHY

1 "Equivalent Linearization, Nonlinear Circuit Synthesis, and the Stabilization and Optimization of Control Systems," by R. W. Bass, *Proceedings of the Symposium on Nonlinear Circuit Analysis*, Brooklyn Polytechnic Institute, October, 1956.

2 Progress Report No. 6, Contract DA-36-034-ORD-1273RD, 1–31 October, 1953.

3 Cf. the author's footnote 4.

4 Progress Report No. 11, Contract DA-36-034-ORD-1273RD, 1–31 March, 1954.

5 "Second Order Contactor Servomechanisms," by H. S. McDonald, Thesis, The Johns Hopkins University, June, 1955.

6 "The Analysis and Synthesis of Relay and Nonlinear Servosystems," by R. W. Bass. [Copyrighted June 30, 1955 by R. W. Bass, this report was distributed as Section III of the Final Report of The Johns Hopkins University Institute for Cooperative Research on the contract mentioned in footnote 2.]

7 "Improved On-Off Missile Stabilization," by R. W. Bass, *Jet Propulsion*, August, 1956.

8 "A Generalization of the Functional Relation $Y(t + s) = Y(t)Y(s)$ to Piecewise-Linear Difference-Differential Equations," by R. W. Bass, to appear in the next issue of *Quarterly of Applied Mathematics*.

Author's Closure

It is interesting to know that Dr. Bass and others have been working along the same lines as the author and it will be most useful when the material quoted by Dr. Bass is made generally available. The author's own interest in the technique began in the summer of 1953 and in December of that year the main outlines of the method were given in a talk at the University of Minnesota, Department of Electrical Engineering. A description of the method with preliminary results was also circulated at the Cambridge University Symposium on Nonlinear Control Problems, held in September, 1954.

Just as it became clear that a method devised for the study of relay control systems could be applied to any piecewise linear system, so has it since become apparent that the method can be applied even more widely to systems with continuously variable gain characteristics and possibly also to those with parameters that are functions of the dependent variable and its derivatives. For such an extension to be admissible the important feature of the response is that it should have a fundamental oscillatory mode that predominates over all others, except possibly a subsidence mode. Under such circumstances the response will correspond to that of a second-order system. This is an extension of the technique of analysis already adopted for the study of linear systems. In practice, many nonlinear systems, as for example an aircraft or missile in flight, are fortunately of the necessary type. Having accepted this approach a useful appli-

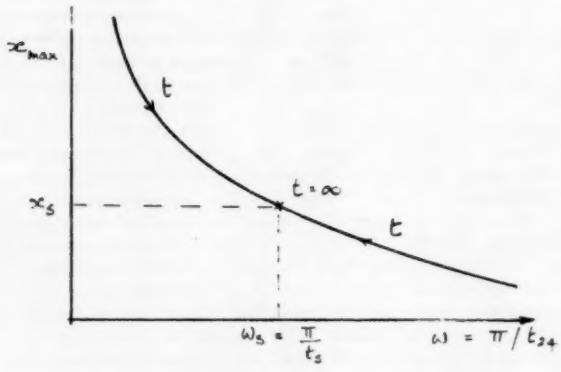


FIG. 32

cation of the method is to the analysis and recording of the empirical information obtained from tests, as observed by Dr. McDonald. This may well yet prove to be the most valuable contribution that this technique has to offer, as there is currently no serious competitor in this field. Characteristic curves derived empirically provide information about system behavior which can be used more directly than frequency-response curves and describing functions, even if these are obtainable for a system functioning as a whole.

It may be noted also that a typical transient response can be represented by a curve similar to that shown in Fig. 32, which shows the variation in amplitude with frequency and can be used to obtain immediately the response from any initial amplitude. Fig. 32 shows a system with a stable oscillation of amplitude x_s and frequency ω_s.

It will be appreciated from the foregoing remarks that the virtual restriction of this method to second-order systems is not so severe a limitation as might at first appear, and that furthermore the essential simplicity thus obtained makes it applicable not only to analysis but also, in some cases, to problems of synthesis and optimization. In particular, it affords a particularly convenient approach for the treatment of linear systems with dead band and hysteresis, as the characteristic curves of any linear system must be straight lines.

On the Dynamics of Pneumatic Transmission Lines

By C. P. ROHMANN[1] AND E. C. GROGAN,[2] PHILADELPHIA, PA.

An analysis of pneumatic transmission tubing dynamics based on a system of evenly distributed fixed parameters is presented. The driving-point impedance and the output-input frequency response of pneumatic tubing are derived from the frequency-response approach with small signals. Both calculated and experimental results are plotted for selected lengths of tubing. The correlation between calculated and test results for frequency response with amplitudes of ±5 per cent of conventional full-scale signal (3–15 psig) or less is good. The relations obtained are very useful in determining the effects of output loads on pneumatic-device dynamics and the limitations imposed on process control by various-length pneumatic transmission lines.

Nomenclature

The following nomenclature is used in the paper:

Where upper and lower-case letters are used for the same variable the lower case indicates instantaneous value, the capital letter the vector notation.

Subscript r denotes receiver.

Subscript i denotes input.

Where dimensions are specified, the inch-pound-second system is used.

C = capacitance per unit length = $\dfrac{\text{in.}^5}{\text{lb}} \Big/ \text{in.}$

f_f = shear force, lb

f_m = inertia force, lb

f_p = pressure force, lb

I, i = volumetric flow rate, in.^{3}/sec

j = $\sqrt{-1}$

l = length of tube, in.

L = inertance per unit length, $\dfrac{\text{lb-sec}^2}{\text{in.}^5} \Big/ \text{in.}$

n = polytropic exponent

P, p = pressure deviation from mean, psi

P_0 = mean pressure, psia

r = tube inside radius, in.

R = resistance per unit length, $\dfrac{\text{lb-sec}}{\text{in.}^5} \Big/ \text{in.}$

S, s = velocity, ips

t = time, sec

v = specific volume, in.^{3}/lb

x = distance from center of tube, in.

X_L = inertant reactance per unit length, $\dfrac{\text{lb-sec}}{\text{in.}^4} \Big/ \text{in.}$

Y = shunt admittance per unit length, $\dfrac{\text{in.}^5}{\text{lb-sec}} \Big/ \text{in.}$

Z = impedance per unit length, $\dfrac{\text{lb-sec}}{\text{in.}^5} \Big/ \text{in.}$

Z_0 = characteristic impedance, $\dfrac{\text{lb-sec}}{\text{in.}^5}$

$\alpha + j\beta$ = propagation constants, in.$^{-1}$

δ = mass density, $\dfrac{\text{lb-sec}^2}{\text{in.}^4}$

λ = dimensionless vector quantity defined by Equation [3]

μ = viscosity, $\dfrac{\text{lb-sec}}{\text{in.}^2}$

ω = angular velocity, radians/sec
= $2\pi f$, where f is frequency of oscillation, cycles/sec (cps)

Introduction

Any analysis of the dynamics of pneumatic control systems must include the transmission tubing. Since the tubing imposes a load on the pneumatic devices, it affects device dynamics as well as the relation between device output and the signal received at the end of the transmission lines. Since there is this interrelation between the device and the transmission line, it is desirable to define the driving-point impedance of tubing in terms that will permit the use of modern operational methods of analysis.

Experimental data on pneumatic transmission have been published by Eckman and Gess (1)[3] and by Bradner (2). Iberall (3) has presented a theoretical analysis of the input-output relation for a tube terminated in a volume subject to a sinusoidal pressure signal at the input. This analysis is based on incompressible viscous flow in the tube and is then modified to account for compressible flow in the tube, finite-signal amplitude, end effects, fluid inertia, and heat transfer. Moise (4) has published a linearized solution for the frequency response of volume-terminated tubing. He also presented experimental data on the frequency response of long lengths of tubing when driven by sinusoidal pressure signals of different amplitudes.

The analysis presented in this paper is linearized to a form selected for convenience in engineering computation of the dynamics of complete pneumatic systems. Vector notation is used and the pneumatic parameters of resistance, capacitance, and inertance introduced, corresponding to the analogous electrical quantities of resistance, capacitance, and inductance. Pressure is made analogous to voltage, and volumetric flow rate corresponds to current. By this approach the considerable body of information on electric transmission lines now in the literature is made available for analysis of pneumatic transmission! In order to make this form applicable the following assumptions and limitation must be imposed: The pressure-signal amplitude is small relative to the total absolute pressure. The end effects are negligible. There is no steady flow through the tube. The final as-

[1] Research Engineer, Minneapolis-Honeywell Regulator Company, Brown Instruments Division. Assoc. Mem. ASME.

[2] Research Engineer, Minneapolis-Honeywell Regulator Company, Brown Instruments Division.

Contributed by the Dynamic Systems Committee of the Instruments and Regulators Division and presented at a joint session with the Machine Design Division at the Semi-Annual Meeting, Cleveland, Ohio, June 17–21, 1956, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, December 13, 1955. Paper No. 56—SA-1.

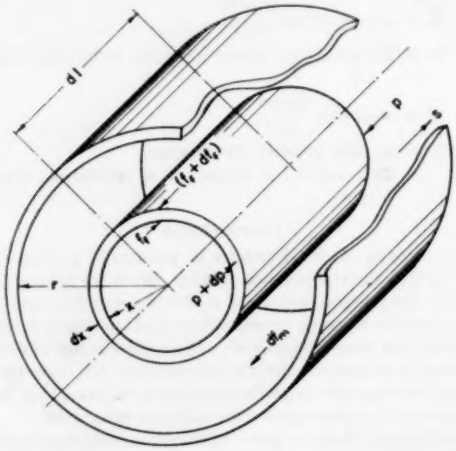
[3] Numbers in parentheses refer to the Bibliography at the end of the paper.

sumption is that the resistance, inertance, and capacitance parameters are distributed uniformly throughout the tube.

Plots are included in this paper showing the correlation between the derived expressions and actual test results. The tests were conducted using $^3/_{16}$-in-ID standard $^1/_4$-in-OD drawn copper tubing with selected terminal volumes, since this is the most commonly used tubing. One test with $^3/_8$-in-OD copper tubing is presented for comparison. A mean level of 9 psig was selected for the tests as this is mid-scale of the standard industrial pneumatic-control range. Emphasis is placed on driving-point impedance and short runs where resonances occur, since test data on these aspects of transmission have not been previously published.

Derivation of Relations Governing Frequency Response of Tubing

Development of Resistance and Inertance Parameters for Tubing. The resistance parameter R arises from the effects of viscous drag between neighboring fluid particles. Referring to Fig. 1 and



s = velocity
p = pressure
f_f = friction force
f_m = inertia force

Fig. 1

reviewing briefly, when flow is laminar the net viscous shear force df_f acting on the surfaces of an incremental annular ring is

$$df_f = (f_f + df_f) - f_f = 2\pi(x + dx)dl\mu\left(\frac{\partial s}{\partial x} + \frac{\partial^2 s}{\partial x^2}\,dx\right) - 2\pi x\,dl\mu\,\frac{\partial s}{\partial x} = 2\pi\,dl\mu\left(x\,\frac{\partial^2 s}{\partial x^2} + \frac{\partial s}{\partial x}\right)dx$$

where

x = distance from center of tube
μ = fluid viscosity
$\frac{\partial s}{\partial x}$ = velocity gradient
$2\pi x\,dl$ = inner surface of increment
$2\pi(x + dx)dl$ = outer surface of increment

Expressing the mass of the fluid increment as $2\pi x\delta\,dx\,dl$ (where δ is the mass density) the inertia force will be

$$df_m = -(2\pi x\delta\,dx\,dl)\,\frac{\partial s}{\partial t}$$

where t = time.

Setting the sum of the friction, inertia, and pressure force, ($df_p = 2\pi x\,dx\,dp$), equal to zero and simplifying

$$2\pi x\,dx\,dp + 2\pi\mu dl\left(x\,\frac{\partial^2 s}{\partial x^2} + \frac{\partial s}{\partial x}\right)dx - 2\pi x\,dx\,dl\delta\,\frac{\partial s}{\partial t} = 0$$

$$\frac{1}{\mu}\,\frac{dp}{dl} = -\frac{\partial^2 s}{\partial x^2} - \frac{1}{x}\,\frac{\partial s}{\partial x} + \frac{\delta}{\mu}\,\frac{\partial s}{\partial t} \qquad \text{.....[1]}$$

In this analysis it is assumed that the pressure is uniform over any plane normal to the tube axis. At low frequencies of oscillation when the fluid inertia forces are negligible, (i.e., when $\frac{\delta}{\mu}\,\frac{\partial s}{\partial t}$ is insignificant) the solution for velocity as a function of distance from the center of the tube is

$$s = \frac{dp}{dl}\left[\frac{r^2 - x^2}{4\mu}\right]$$

where r = tube internal radius.

The velocity distribution across the tube at any instant of time is parabolic with zero velocity at the tube wall. Expressing the volumetric flow rate through the tube i, as a function of the pressure gradient, $\frac{dp}{dl}$

$$i = \int_0^r 2\pi x s\,dx$$

$$i = \frac{\pi r^4}{8\mu}\,\frac{dp}{dl}$$

or $\frac{dp}{dl} = iR$, where $R = \frac{8\mu}{\pi r^4}$, the resistance per unit length of the tube.

The inertance parameter for tubing results from the inertia property of the air and becomes significant at high fluid acceleration, which occurs at high frequencies of oscillation. In the case where the acceleration forces dominate over viscous-drag effects, the velocity becomes uniform across the tube and Equation [1] reduces to

$$\frac{dp}{dl} = \delta\,\frac{\partial s}{\partial t}$$

Confining the discussion to flow and pressure signals which vary sinusoidally with time and using vector notation, the imaginary quantity $j\omega$ may be substituted for the $\partial/\partial t$ operator

$$\frac{dp}{dl} = j\omega\delta S$$

where $\omega = 2\pi f$, with f the frequency of oscillation, expressed in cycles per second (cps).

Again in this case, relating the pressure gradient to the flow across the section

$$I = \int_0^r 2\pi S x\,dx = \frac{\pi r^2}{j\omega\delta}\,\frac{dp}{dl}$$

or

$$\frac{dp}{dl} = IX_L$$

where

$$X_L = j\omega L$$

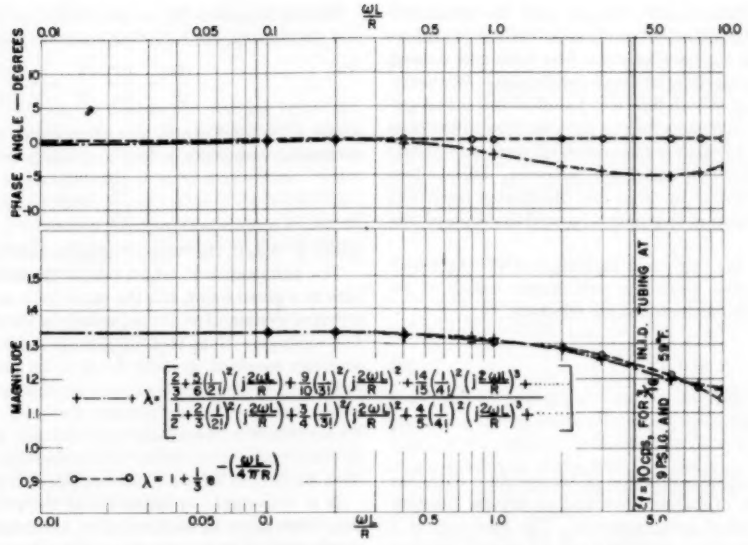


FIG. 2 CALCULATED CURVES OF APPROXIMATE VERSUS EXACT SOLUTIONS OF λ

and

$$L = \frac{\delta}{\pi r^2}$$

the inertance of the air in the tube per unit length.

Over intermediate frequency ranges where both inertia and viscous-drag effects are significant in governing fluid motion, the pressure gradient becomes the following function of the net flow through a given section of the tube

$$\frac{dp}{dl} = I(R + \lambda X_L) \quad \ldots\ldots\ldots\ldots\ldots [2]$$

where

$$R = \frac{8\mu}{\pi r^4}; \quad X_L = j\omega L = \frac{j\omega\delta}{\pi r^2}$$

and

$$\lambda = \left[\frac{\frac{2}{3} + \frac{5}{6}\left(\frac{1}{2!}\right)^2\left(\frac{2X_L}{R}\right) + \frac{9}{10}\left(\frac{1}{3!}\right)^2\left(\frac{2X_L}{R}\right)^2 + \frac{14}{15}\left(\frac{1}{4!}\right)^2\left(\frac{2X_L}{R}\right)^3 + \ldots}{\frac{1}{2} + \frac{2}{3}\left(\frac{1}{2!}\right)^2\left(\frac{2X_L}{R}\right) + \frac{3}{4}\left(\frac{1}{3!}\right)^2\left(\frac{2X_L}{R}\right)^2 + \frac{4}{5}\left(\frac{1}{4!}\right)^2\left(\frac{2X_L}{R}\right)^3 + \ldots}\right] \quad \ldots\ldots [3]$$

this function can be approximated by

$$\lambda = 1 + \frac{1}{3}e^{-\left(\frac{\omega L}{4\pi R}\right)} \quad \ldots\ldots\ldots\ldots [4]$$

as shown in Fig. 2.

The general solution of Equation [1] is given in the Appendix.

The impedance quantity $(R + \lambda X_L)dl$, which relates pressure drop to flow over an incremental length of tubing is thus seen to be more complex than that of the simple *R-L* arrangement used in the electrical transmission-line analogy, owing to the changing nature of the velocity profile with frequency. It is interesting to note that

$$\lim_{\omega \to 0} (R + \lambda X_L) = \left(R + \frac{4}{3} X_L\right)$$

Thus, at low frequencies, where the velocity distribution is parabolic, the effective inductance of tubing per unit length is 4/3 the value it assumes at high frequencies.

In the cases tested in this investigation, however, it was found adequate to assume λ equals unity for all cases. This gives the following equation which is completely analogous to the electric transmission line

$$\frac{dP}{dl} \approx I(R + X_L) \quad \ldots\ldots\ldots\ldots\ldots [5]$$

This approximation causes no error at the two ends of the frequency range. At low frequency the inertant reactance (pneumatic equivalent of inductive reactance) is negligible relative to the resistance. At high frequency the velocity profile flattens out and $\lambda \to 1$ as used in Equation [5]. There is, however, a range of frequencies around $\omega L/R = 1$ through which the magnitude of the per unit length impedance given in Equation [2] could be 20 per cent higher than that given by Equation [5]. Also the phase angle associated with the former could be 7 deg larger than that given by Equation [5]. The preceding work is all based on laminar flow, however. Turbulence tends to flatten the velocity profile and hence reduce or eliminate the error in the effective inertance. In no case in the test work has the correlation between test and calculated relations been observed to be adversely affected by use of the approximate Relation [5].

The resistance derived in the foregoing work is based on fully developed laminar flow. Turbulence would tend to increase the resistance. The mean per unit length resistance was derived from dynamic test data and was found to vary with tubing length and terminating conditions as well as with the amplitude and frequency of the drive signal which determine flow amplitude. In general, for 1.2 psi peak-to-peak drive signals over intermediate frequencies, and where tubing lengths exceed 100 ft, the effective

resistance was found to exceed by 20 per cent the calculated value based on static laminar-flow conditions. Moise (4) calculated friction factors on the basis of static flow tests and showed improved correlation, especially at large amplitudes. While the same trend of increasing R with flow was found in this investigation, it was considered sufficient for the purpose here to increase R by 20 per cent over the calculated value in all cases considered. Where turbulence becomes excessive, however, friction-loss effects cannot be represented by an evenly distributed resistance parameter. The treatment of that case is beyond the scope of this paper.

Summarizing then, for the cases investigated it was found suitable to determine the mean per unit length values of resistance and inertance from the following relations

$$R = 1.2\left(\frac{8\mu}{\pi r^4}\right) \quad [6]$$

$$L = \frac{\delta}{\pi r^2} \quad [7]$$

Since the viscosity of air is a function of temperature only, the resistance is a function of temperature and an inverse function of the tube cross-sectional area, squared. The inertance is a direct function of the absolute pressure level, an inverse function of absolute temperature, and an inverse function of the cross-sectional area of the tube.

Development of Capacitive Parameter for Tubing. The capacitive parameter for tubing results from the compressibility of air. In the development of expressions for the pressure gradient, the effects of possible incremental changes in the net volumetric flow rate of air particles passing through an incremental length of tubing dl were ignored, since

$$\lim_{dl \to 0} (i + di) = i$$

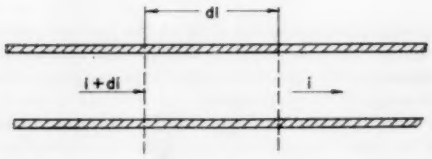


FIG. 3

However, at a given point in the tube, pressure variations with time may be related directly to the difference between volumetric flow rates across an incremental length of tube. Consider an increment of tube of length dl with inflow $(i + di)$ and outflow i, Fig. 3. The net inflow into the section is

$$di = -\frac{\pi r^2 dl}{v}\frac{dv}{dt} \quad [8]$$

where

v = specific volume of air in section
r = radius of tube
i = volumetric flow rate

Assuming a polytropic compression with the resultant relation $pv^n = \text{const}$, and differentiating with respect to time

$$npv^{(n-1)}\frac{dv}{dt} + v^n\frac{dp}{dt} = 0 \quad [9]$$

where p = pressure, absolute units, and n = polytropic exponent.

Solving Equation [8] for $(dv)/(dt)$, substituting in Equation [9], and simplifying, the following equation is obtained

$$\frac{di}{dl} = \frac{\pi r^2}{np}\frac{dp}{dt} = C\frac{dp}{dt} \quad [10]$$

where $C = (\pi r^2)/(np)$, the capacitance per unit length. For sinusoidal variations in flow and using vector notation

$$\frac{dI}{dl} = YP \quad [11]$$

where $Y = j\omega C$, the "shunt" admittance of tubing per unit length.

The admittance Y, which relates the flow gradient to the pressure at a given point, has the same form as that of an evenly distributed system of shunt capacitors in the electrical transmission-line analogy. Note that the absolute pressure p is assumed essentially constant in order for C to be constant. This then restricts the derived expression to small amplitude pressure signals about a constant mean pressure level, i.e., where tubing is either "dead ended" or terminated in a volume, and the d-c flow is zero at every point in the tube. The tests reported in this paper have been limited to 1.2 psi peak-to-peak and smaller amplitudes.

It is reasonable to assume that the polytropic exponent will vary from unity or isothermal, at low frequencies up toward 1.4, corresponding to a reversible adiabatic compression, as the frequency increases, owing to reduced heat-transfer periods. For $^3/_{16}$-in-ID tubing an assumption of $n = 1$ has proven satisfactory up to several cycles per sec.

Referring back to Equation [10] it can be seen that the capacitance per unit length of tubing varies directly with the cross-sectional area of the tube and inversely with the mean pressure level. Additional capacitance due to changing tube volume with varying pressure is entirely negligible for commercial metallic tubing and standard operating levels.

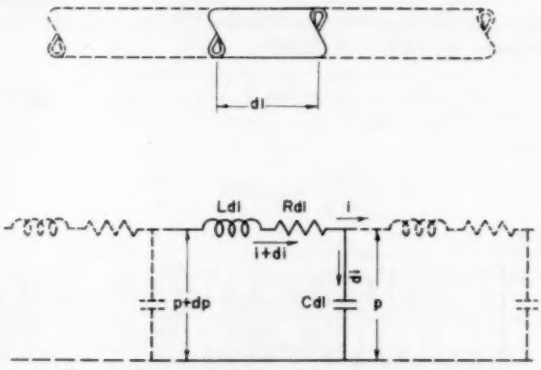


FIG. 4 PNEUMATIC LINE AND ELECTRICAL ANALOG

Development of Transmission-Line Equations for Tubing. In the foregoing it was demonstrated that with certain reservations, the expressions for the flow and pressure gradient permitted the effects of fluid inertia, viscosity, and compressibility to be presented by a fixed system of evenly distributed L-R-C parameters, analogous to the electric transmission line as shown in Fig. 4. The equations for flow (I) and pressure (p) developed at any distance l from the receiving end, and at a given frequency, may be found by combining Equations [5] and [11] and using the boundary condition that for $l = 0$; $P = P_r$; $I = I_r$

$$P = \frac{1}{2}\left[\left(P_r + I_r\sqrt{\frac{Z}{Y}}\right)e^{\sqrt{ZY}\,l} + \left(P_r - I_r\sqrt{\frac{Z}{Y}}\right)e^{-\sqrt{ZY}\,l}\right] \quad [12]$$

$$I = \frac{1}{2}\left[\left(I_r + P_r\sqrt{\frac{Y}{Z}}\right)e^{\sqrt{ZY}\,l} + \left(I_r - P_r\sqrt{\frac{Y}{Z}}\right)e^{-\sqrt{ZY}\,l}\right] \quad \ldots\ldots [13]$$

Equations [12] and [13] may be used as the working equations for the exact solution of long lengths of transmission tubing where the end effects may be neglected. P and I are the vector representation of pressure and volumetric flow at a distance l from the receiver. For the usual case of interest, that of the relation between the pressures at the sending and receiving ends of a tube of total length l, the following equation is obtained by dividing Equation [12] by P_r and inverting

$$\frac{P_r}{P_s} = \frac{2}{[e^{\sqrt{ZY}\,l} + e^{-\sqrt{ZY}\,l}] + \dfrac{I_r}{p_r}\sqrt{\dfrac{Z}{Y}}[e^{\sqrt{ZY}\,l} - e^{-\sqrt{ZY}\,l}]} \quad \ldots\ldots [14]$$

Term $\sqrt{ZY}$ is a complex quantity and may be represented by $(\alpha + j\beta)$ called the "propagation constant" which determines how a pressure wave is propagated through a length of tubing with reference to change in magnitude and phase.

Term $\sqrt{Z/Y}$ has the units of an impedance and in electrical systems is referred to as the surge, or characteristic impedance Z_0 of a line.

As in the case of electrical transmission lines, then, the behavior of transmission tubing is determined completely by two parameters; the characteristic impedance Z_0 and the propagation function $(\alpha + j\beta)l$. A direct method of determining the values of these parameters, on the basis of tests on a long line, will be discussed later in the text.

The relation between P_r and I_r is governed by the impedance of the receiver Z_r and may be expressed as

$$\frac{P_r}{I_r} = Z_r = \frac{1}{Y_r}$$

where Y_r is called the admittance. Z_r or its reciprocal Y_r gives the vector relation between the sinusoidal pressure and flow at the receiver. For example, if the line is terminated in a fixed-volume tank, the flow into the tank is related to the pressure by the vector relation

$$\frac{P_r}{I_r} = Z_r = \frac{1}{j\omega C_r} \quad \ldots\ldots [15]$$

where

C_r is capacitance of tank $= \dfrac{V_r}{nP_0}$

V_r = tank volume
n = polytropic exponent
P_0 = mean pressure level in absolute units

Equation [14] may be revised so that the input-output pressure relation or transfer function may be expressed as a function of the admittance of the load and the characteristic impedance and propagation constants of the tubing as follows

$$\frac{P_r}{P_s} = \frac{2}{[e^{(\alpha+j\beta)l} + e^{-(\alpha+j\beta)l}] + Y_rZ_0[e^{(\alpha+j\beta)l} - e^{-(\alpha+j\beta)l}]} \quad [16]$$

Similarly the driving-point impedance of tubing Z_s, which relates the net rate of flow I_s into the sending end to the input driving pressure P_s, may be derived by dividing Equation [12] by Equation [13], yielding

$$Z_s = \frac{P_s}{I_s} = Z_0 \left\{\frac{[e^{(\alpha+j\beta)l} + e^{-(\alpha+j\beta)l}] + Y_rZ_0[e^{(\alpha+j\beta)l} - e^{-(\alpha+j\beta)l}]}{Y_rZ_0[e^{(\alpha+j\beta)l} + e^{-(\alpha+j\beta)l}] + [e^{(\alpha+j\beta)l} - e^{-(\alpha+j\beta)l}]}\right\} \quad \ldots\ldots [17]$$

Equation [17] which defines the loading effects of tubing is of particular use to those concerned with the evaluation of the dynamic performance of pneumatic devices with varied output loads. The dependency of the driving-point impedance on tubing length and the load in which the tube is terminated can be seen from inspection of this equation.

Through the use of hyperbolic functions, Equations [16] and [17] may be reduced to the more simple forms

$$\frac{P_r}{P_s} = \frac{\sinh\,[A + jB]}{\sinh\,[(A + \alpha l) + j(B + \beta l)]} \quad \ldots\ldots [18]$$

$$Z_s = Z_0 \tanh\,[(A + \alpha l) + j(B + \beta l)] \quad \ldots\ldots [19]$$

where the quantity $[A + jB]$ may be determined from

$$\tanh\,[A + jB] = \frac{Z_r}{Z_0} \quad \ldots\ldots [20]$$

The propagation constants α and β, and the characteristic impedance Z_0, may be determined from the expressions developed earlier in this paper or from the summary of relations given in the Appendix. For $^3/_{16}$-in-ID tubing at 59 F and 9 psig mean level, the values of these functions are given in Figs. 5 and 6, which cover a frequency range of 2 to 600 cpm.

In evaluating Equations [18] and [19] and the function $[A + jB]$, it will be necessary to refer to charts of hyperbolic functions of a complex variable. Such charts have been published by Kimbark (10), Kennelly (11), Smith (12), and others. Figs. 7 and 8 are charts of $y = \tanh x$ and $y = \sinh x$, respectively, published by Kimbark. In these charts y is located as a complex variable in the rectangular co-ordinates, and x in the curvilinear co-ordinates.

Lumped System Approximations. In analog-computer studies, it may be necessary to develop lumped approximations to the distributed system. In these studies, the expanded forms of Equations [16] and [17] are used. For example, in the case of a tubing run terminated in a volume, and representing $e^{\pm(\alpha+j\beta)l}$ by a power series, Equations [16] and [17] become the following convergent series expressions

$$\frac{P_r}{P_s} = \frac{1}{1 + \left(\dfrac{2g_0+1}{2!}\right)\left[RCl^2j\omega\left(\dfrac{L}{R}j\omega + 1\right)\right] + \left(\dfrac{4g_0+1}{4!}\right)\left[RCl^2j\omega\left(\dfrac{L}{R}j\omega + 1\right)\right]^2 + \ldots} \quad \ldots\ldots [21]$$

$$Z_s = \frac{1}{(1+g_0)Clj\omega}\left\{\frac{1 + \left(\dfrac{2g_0+1}{2!}\right)\left[RCl^2j\omega\left(\dfrac{L}{R}j\omega + 1\right)\right] + \left(\dfrac{4g_0+1}{4!}\right)\left[RCl^2j\omega\left(\dfrac{L}{R}j\omega + 1\right)\right]^2 + \ldots}{1 + \dfrac{1}{3!}\left(\dfrac{3g_0+1}{g_0+1}\right)\left[RCl^2j\omega\left(\dfrac{L}{R}j\omega + 1\right)\right] + \dfrac{1}{5!}\left(\dfrac{5g_0+1}{g_0+1}\right)\left[RCl^2j\omega\left(\dfrac{L}{R}j\omega + 1\right)\right]^2 + \ldots}\right\} \quad \ldots\ldots [22]$$

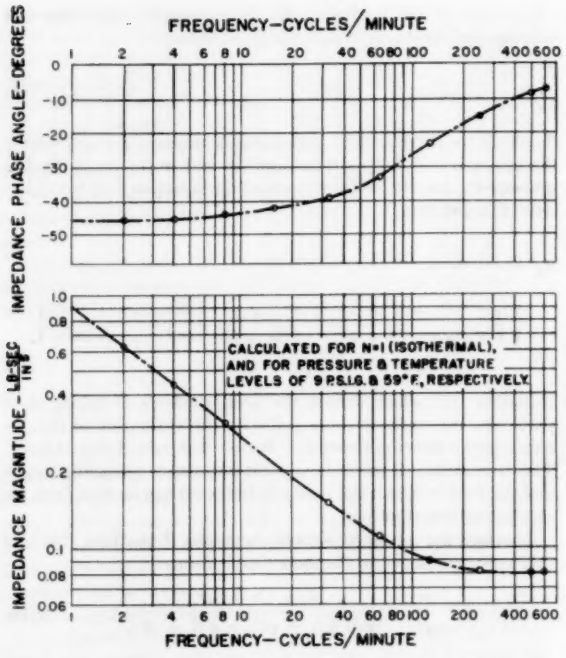


FIG. 5 CHARACTERISTIC IMPEDANCE Z_0 OF 3/16-IN-ID COPPER TUBING

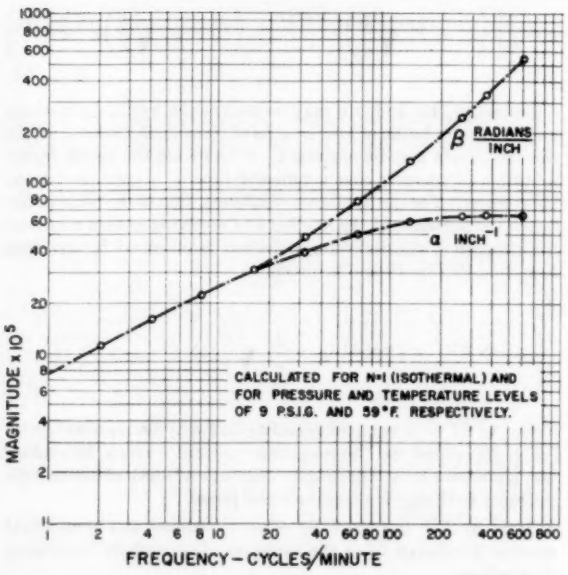


FIG. 6 PROPAGATION CONSTANTS OF 3/16-IN-ID COPPER TUBING

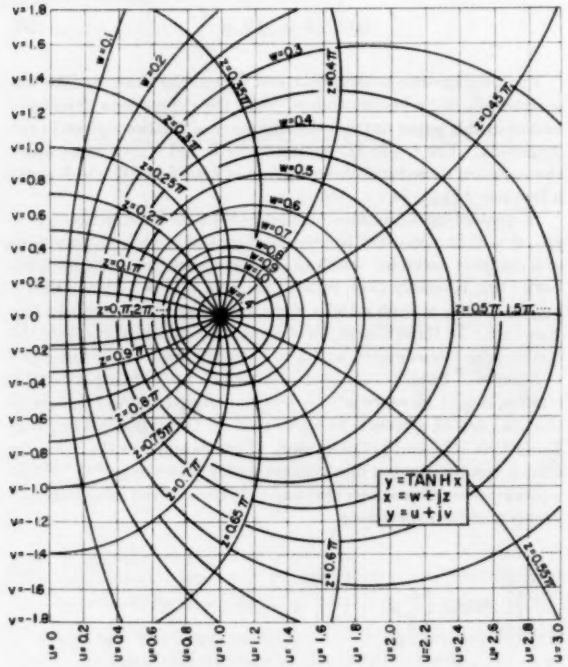


FIG. 7 HYPERBOLIC TANGENT CHART

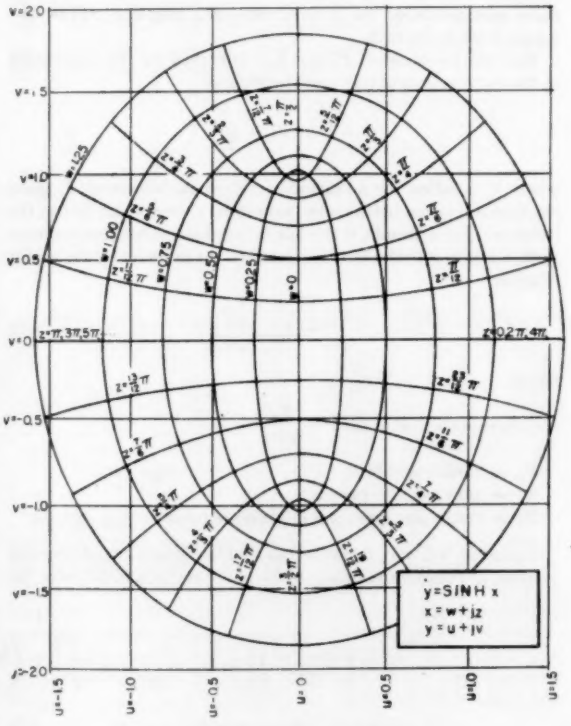


FIG. 8 HYPERBOLIC SINE CHART

where, for C_r = the volumetric capacitance of the receiver

$$g_0 = \frac{C_r}{Cl}$$

Over low-frequency ranges where Equations [21] and [22] may be approximated by

$$\frac{P_r}{P_i} \approx \frac{1}{1 + \left(\frac{2g_0 + 1}{2!}\right)\left[RCl^2 j\omega\left(\frac{L}{R} j\omega + 1\right)\right]}$$

$$Z_i \approx \frac{1}{(1 + g_0)Clj\omega}\left\{\frac{1 + \left(\frac{2g_0 + 1}{2!}\right)\left[RCl^2 j\omega\left(\frac{L}{R} j\omega + 1\right)\right]}{1 + \frac{1}{3!}\left(\frac{3g_0 + 1}{g_0 + 1}\right)\left[RCl^2 j\omega\left(\frac{L}{R} j\omega + 1\right)\right]}\right\}$$

a length of tubing terminated in a volume with capacitance C_r may be represented by the electrical circuit given in Fig. 9 which has identical transfer lag and input impedance properties.

Over broader frequency ranges where the higher-order terms appearing in Equations [21] and [22] become significant, the approximations become more complex. The method of determining the approximations and the error introduced is discussed in the electrical literature such as reference (5).

Test Methods and Results

Tests were run using two sizes of standard drawn copper tubing 1/4 in. OD—0.186 in. ID and 3/8 in. OD—0.291 in. ID. The emphasis was placed on the smaller tubing since this is the most commonly used size. Termination volumes of 1.2 cu in., and 92.5 cu in. were used. Special connection fittings were used to keep the ID constant for long lengths. The tube was smoothly coiled with an 18-in. diam. Fig. 10 is a schematic diagram of the test setup. Reference (7) gives details of the sine-wave generator used. Flow into the tubing was determined by measuring the pressure across a small glass tube for which the flow-pressure-frequency relation had been determined by test. Flow for the tube calibration was determined from the pressure variation in a fixed-volume tank into which the capillary discharged. A polytropic exponent of 1.0 was assumed in the calibration calculation.

The mean pressure was held at 9 psig for all tests. Peak-to-peak drive amplitudes from 1.2 psi down to 0.17 psi were used. The drive amplitude was held constant for most tests as frequency was varied. Pressure was measured with a Statham gage read on an oscilloscope. Pressure calibration was determined at low frequency against an accurate gage. Since the same Statham gage was used for all measurements, calibration accuracy was not important. Phase was determined by means of Lissajous patterns on the oscilloscope. In all cases only the estimated fundamental component was measured and distortion neglected. Distortion of sufficient magnitude to markedly affect the Lissajous figure was encountered only at the highest frequencies tested and most of this was introduced by the input flow measuring tube. The accuracy of measurement attained of course varied with the size of the signal, the frequency, and the distortion encountered. As an indication of the order of magnitude of the reading errors expected, the following average deviations are listed: ±10 deg phase, ±0.05 on amplitude ratios, and ±10 per cent of the impedance magnitude.

Comparison of the Test Results With the Theory. For ease of comparison the test results (Figs. 11 to 14) are plotted on the same sheets with the points calculated on the basis of the theory. Calculated points for both isothermal and reversible adiabatic

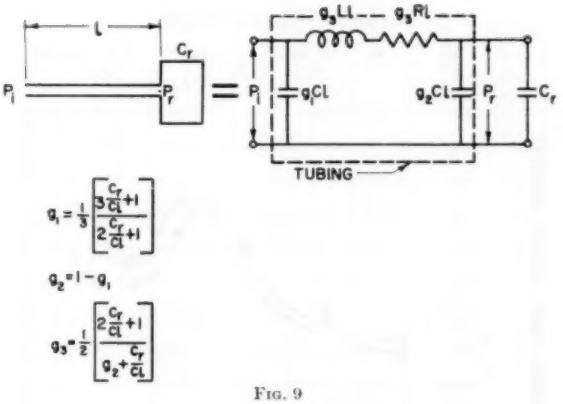


Fig. 9

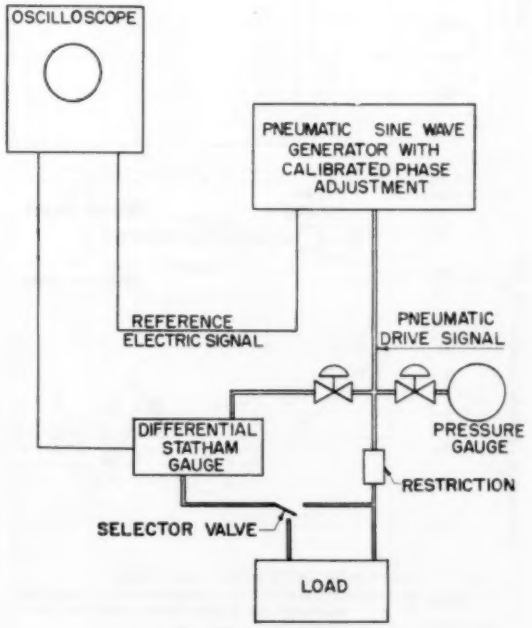


Fig. 10 Test Setup

compression are plotted. The calculations were modified to include the effects of volumetric loading at the gages. As can be seen from the plots, the experimental results tend to follow the isothermal relation over much of the frequency range investigated. For the shorter tubes where resonances occur, impedance changes rapidly with frequency at the higher frequencies. Thus it can be seen that while the trend for the isothermal and adiabatic are not too different, the actual values at any high frequency may be considerably different for the two cases. For this reason the uncertainty in polytropic exponent causes considerable doubt in estimating the impedance at any particular high frequency. Fortunately, the general trend can be estimated reasonably well and this is all that is required for many applications of this theory.

As shown in Figs. 15 and 16 discrepancies between theory and test occur at resonance points where the impedance becomes very small. For the constant drive amplitudes used in the test, the flows become large, causing an increase in the effective resistance

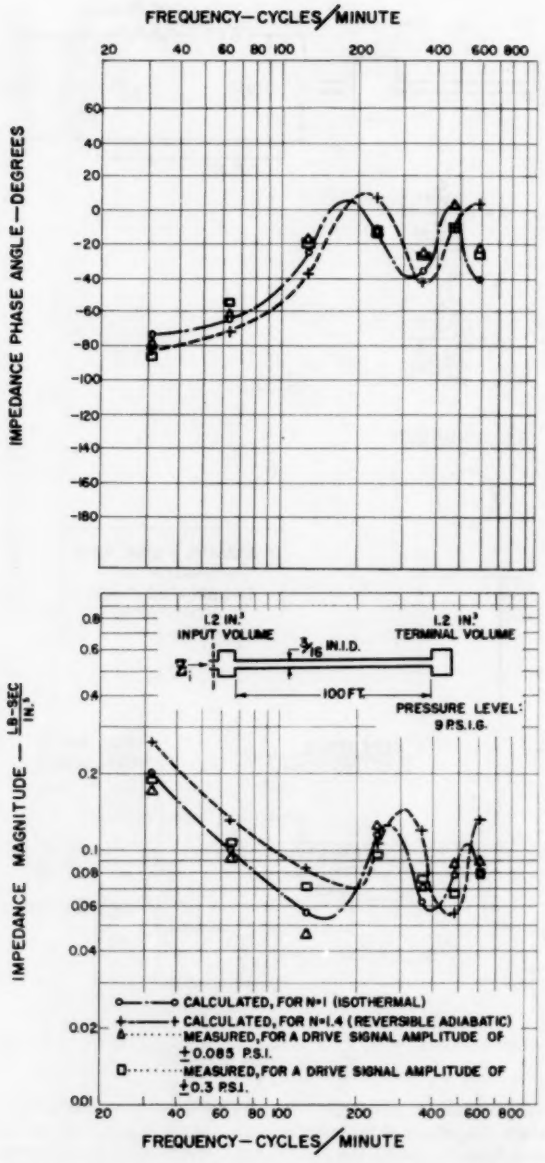


FIG. 11 DRIVING-POINT IMPEDANCE OF 100 FT OF 3/16-IN-ID COPPER TUBING

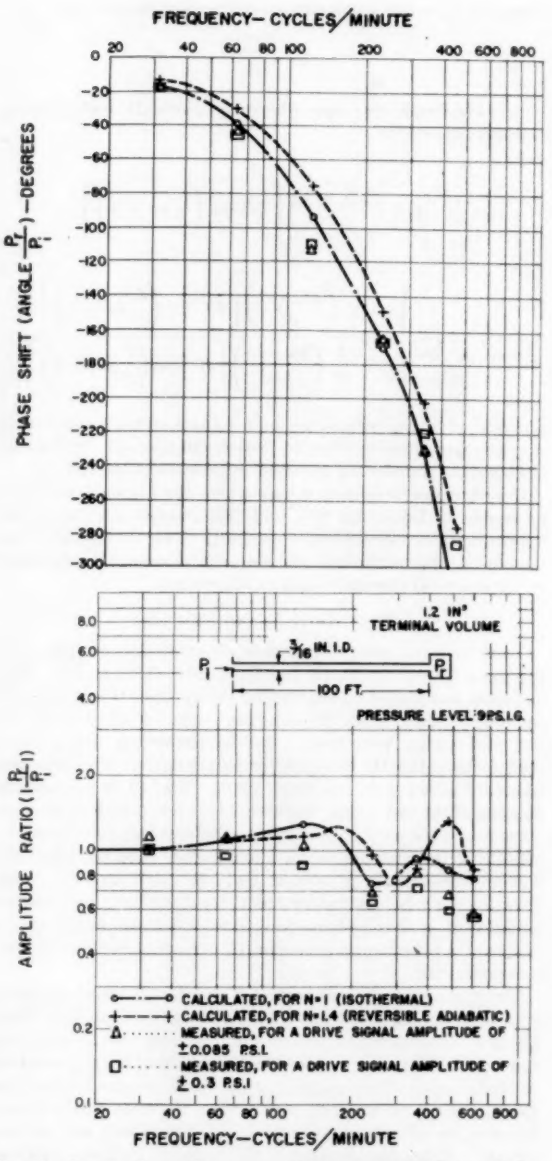


FIG. 12 OUTPUT/INPUT FREQUENCY RESPONSE OF 100 FT OF 3/16-IN-ID COPPER TUBING

parameter over the constant value assumed in the theory. A method of adjusting the resistance with flow might be used to reduce this error. The effect of this increased resistance is to smooth out the curves but not to shift the basic trends. Other than in these two areas, the correlation is quite satisfactory.

Determination of Tubing Parameters From "Long-Line" Tests. Figs. 17 to 19 are the results of tests performed on a 750-ft dead-ended run of 3/16-in-ID transmission tubing. For all practical purposes, the line may be considered to be of infinite length; therefore one without reflections. In these tests, pressure variations were recorded at distances of 30 ft, 80 ft, and 180 ft from the sending end. At these points, the tubing runs from the sending end could be assumed to be terminated in their characteristic impedance Z_0. The response of pressure signals over the distance plotted in Figs. 17 and 18 may then be related to the driving signal P_i simply by

$$\frac{P_l}{P_i} = e^{-(\alpha+j\beta)l} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [23]$$

At a given frequency, the amplitude of the pressure signal will, therefore, be an inverse exponential function of the distance from the sending end. Simultaneously, the phase shift ($-\beta l$) will be a direct function of this distance. The slopes of the lines plotted in Figs. 17 and 18 then offer a quick means of evaluating the propagation constants of the line. The experimentally determined

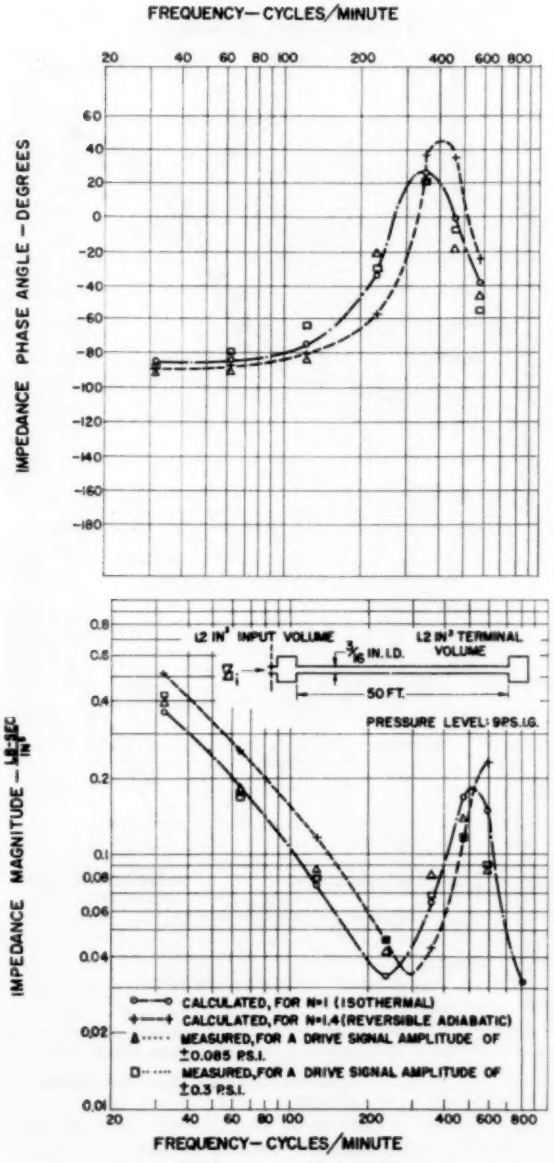


FIG. 13 DRIVING-POINT IMPEDANCE OF 50 FT OF 3/16-IN-ID COPPER TUBING

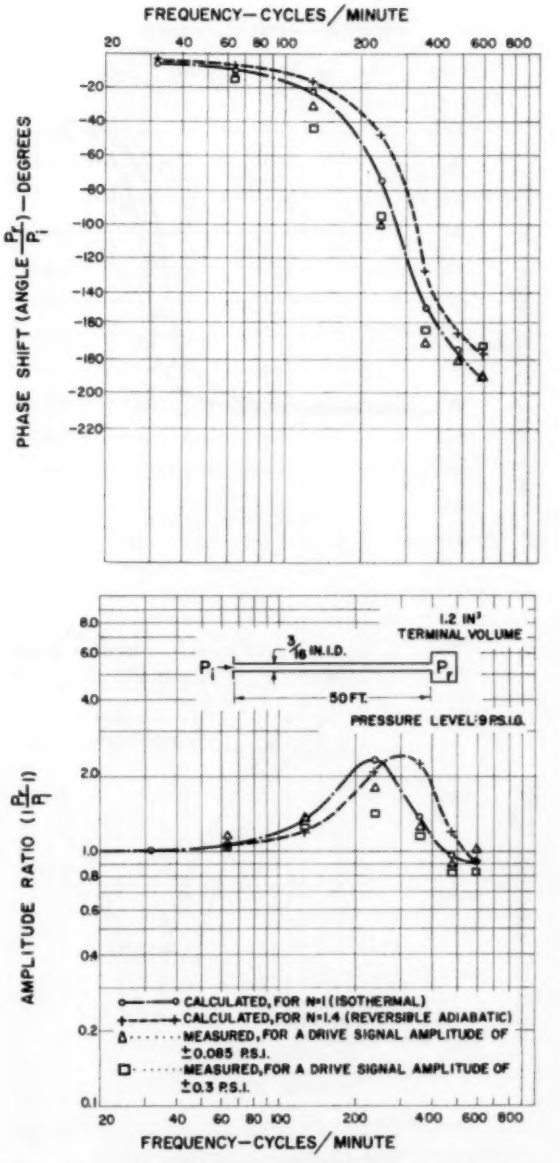


FIG. 14 OUTPUT/INPUT FREQUENCY RESPONSE OF 50 FT OF 3/16-IN-ID COPPER TUBING

values of α and β, given in Figs. 17 and 18, correlate closely with the calculated values plotted in Fig. 6. Fig. 19 is a plot of the driving-point impedance of the 750-ft run, which is shown to correspond with the theoretical characteristic impedance Z_0 of the line.

With the data thus obtained on α, β, and Z_0, the behavior of the line under any operating condition may be determined by reference to Equations [16] and [17], or Equations [18] and [19].

The frequency range covered in the plots is limited to that over which the volume of the input and receiving Statham gages may be neglected.

DISCUSSION AND CONCLUSIONS

On the basis of the correlation between test results and the theory it is felt that the theory constitutes a useful working tool for analyzing pneumatic transmission lines. By the development of a form completely analogous to the electric transmission line, the considerable body of work on electric transmission available in the literature has been made applicable to pneumatic transmission. One aspect of transmission covered in the literature is the concept of wave motion and standing waves. Examination of this concept is helpful in developing a physical understanding of

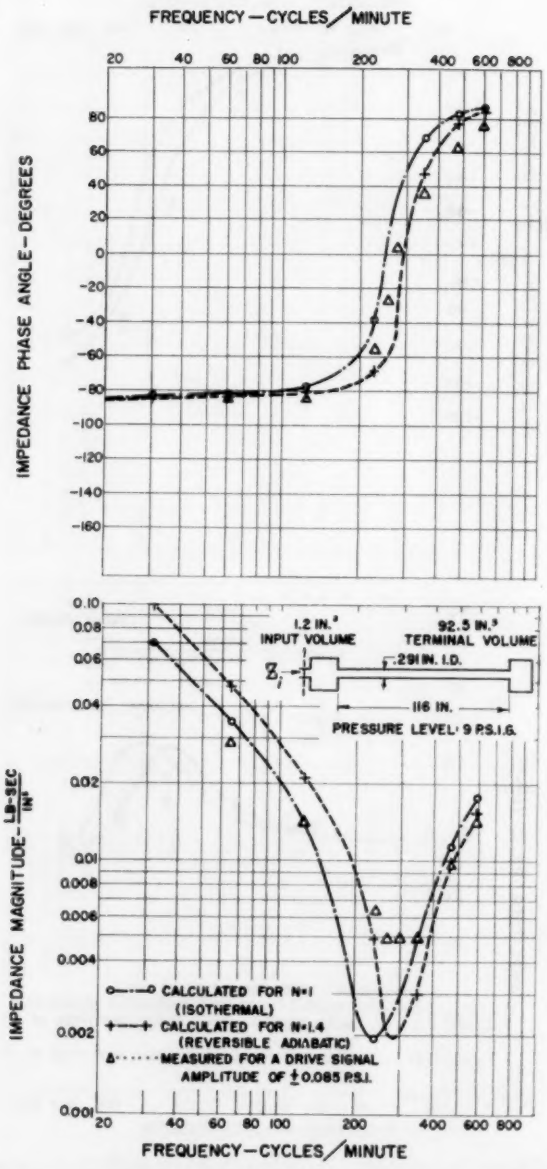


FIG. 15 DRIVING-POINT IMPEDANCE OF 116 IN. OF 0.291-IN-ID COPPER TUBING TERMINATED IN 92.5 CU IN.

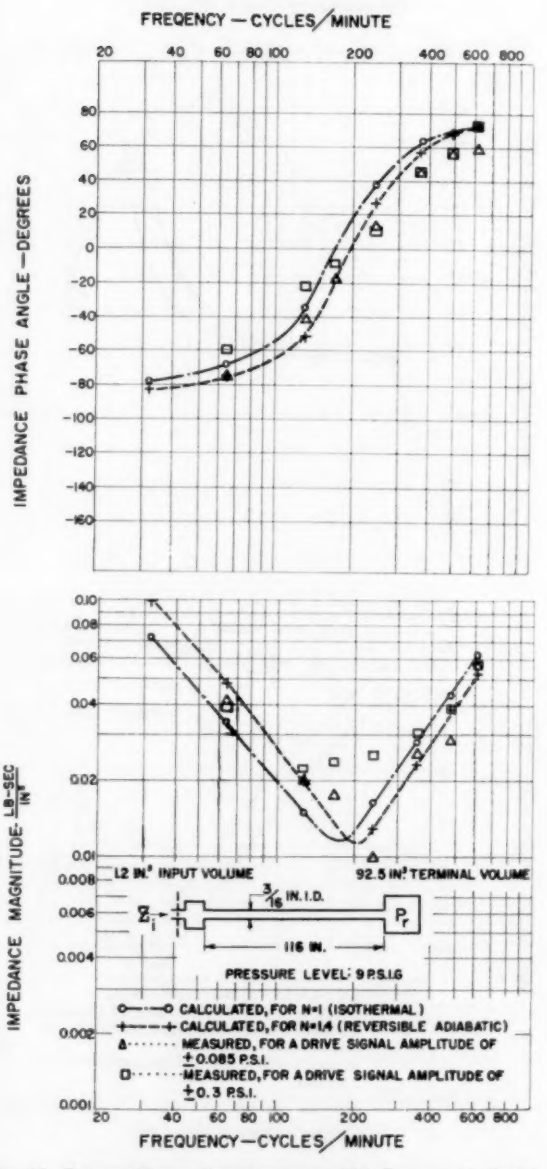


FIG. 16 DRIVING-POINT IMPEDANCE OF 116 IN. OF 3/16-IN-ID COPPER TUBING TERMINATED IN 92.5 CU IN.

certain phenomena, commonly found in acoustics. It is interesting to note that the maximum velocity at which pressure waves may be propagated through a tubing run may be shown to approach as a limit the speed of sound in air.

The velocity of propagation is given by

$$V = \frac{\omega}{\beta}$$

In the Appendix it is shown that

$$\lim_{\omega \to \infty} \beta = \sqrt{LC}\,\omega$$

Thus

$$V_{max} = 1/\sqrt{LC}$$

Substituting the values of L and C developed in this text for pneumatic transmission lines for air at 59 F and, assuming compression to be reversible adiabatic (i.e., $n = 1.4$)

$$LC = 5.6 \times 10^{-9}\ \frac{\text{sec}^2}{\text{in.}^2}$$

$$\frac{1}{\sqrt{LC}} = V_{max} = 1100 \text{ fps}$$

the velocity of sound in air at standard temperature.

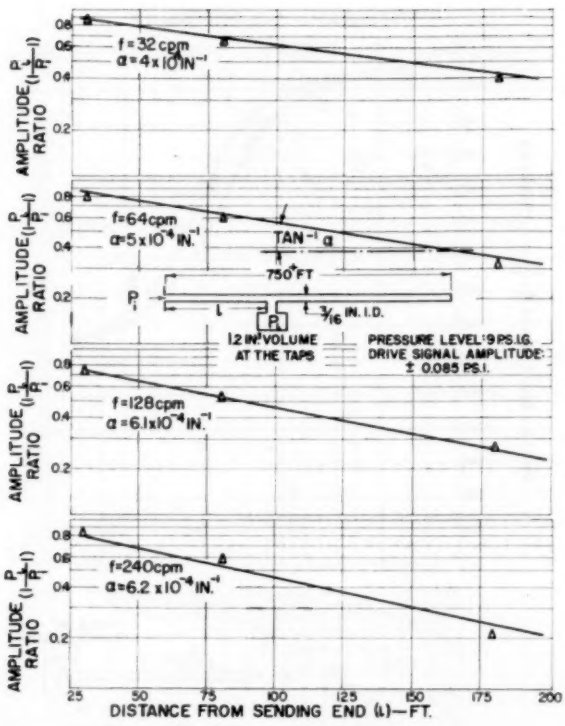


FIG. 17 DETERMINATION OF α FROM LONG-LINE FREQUENCY-RESPONSE DATA

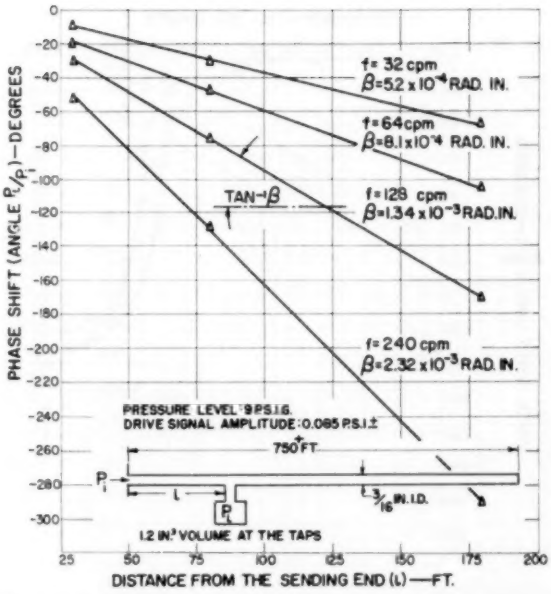


FIG. 18 DETERMINATION OF β FROM LONG-LINE FREQUENCY-RESPONSE DATA

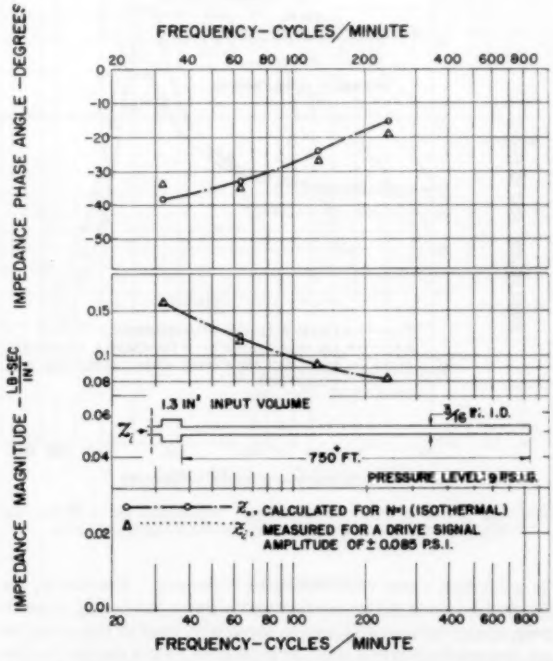


FIG. 19 DETERMINATION OF Z_0 FROM LONG-LINE IMPEDANCE DATA

As clearly shown by the resonances obtained with short lengths of tubing, the inertance or mass effect of air is significant in pneumatic transmisson using common industrial tubing sizes and operating pressure levels. When short runs terminating in large volumes are used, such as shown in Figs. 20 and 21, phase angles exceeding 90 deg can be obtained without attenuation at frequencies on the order of 2 cps. In these cases, the tubing-volume combinations exhibit the properties of underdamped second-order systems. The damping increases as line length increases. For a fixed length, reducing the terminal-load volume or increasing the diameter of the tube reduces the damping but raises the resonant frequency.

In Fig. 14, it is shown that a phase angle of 180 deg was obtained without attenuation for a 50-ft length of $^3/_{16}$-in-ID tubing terminated in a 1-cu-in. volume at a frequency of 7 cps. Fortunately most industrial processes are such that a frequency of this order of magnitude cannot occur. For long lines, Figs. 22 and 23 bring out the advantage of using a valve positioner or volume booster to take the load of the valve top off the transmission line and thereby speed up the transmission.

The loading imposed on a pneumatic device by the output tubing has been shown to be considerably more complex than that of a simple volume. Since this loading has a considerable effect on the device dynamics it is important that it be considered in dynamics specifications. Some reduction of the testing necessary for dynamics evaluation can be realized by examining Equation [17]. Noting that

$$(\alpha + j\beta) = \sqrt{ZY} = j\sqrt{LC}\,\omega\left(1 - j\frac{R}{\omega L}\right)^{1/2}$$

the quantity α, which determines the magnitude of the vectors

$$e^{\pm(\alpha+j\beta)l}$$

will approach

$$\frac{R}{2}\sqrt{\frac{C}{L}}$$

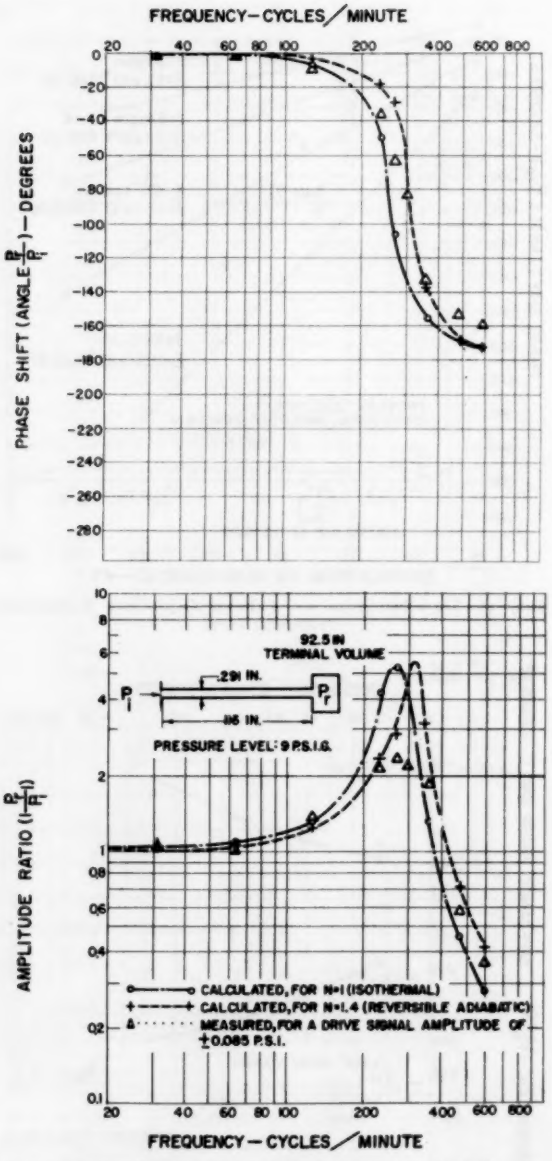


FIG. 20 OUTPUT/INPUT FREQUENCY RESPONSE OF 116 IN. OF 0.291-IN-ID COPPER TUBING TERMINATED IN 92.5 CU IN.

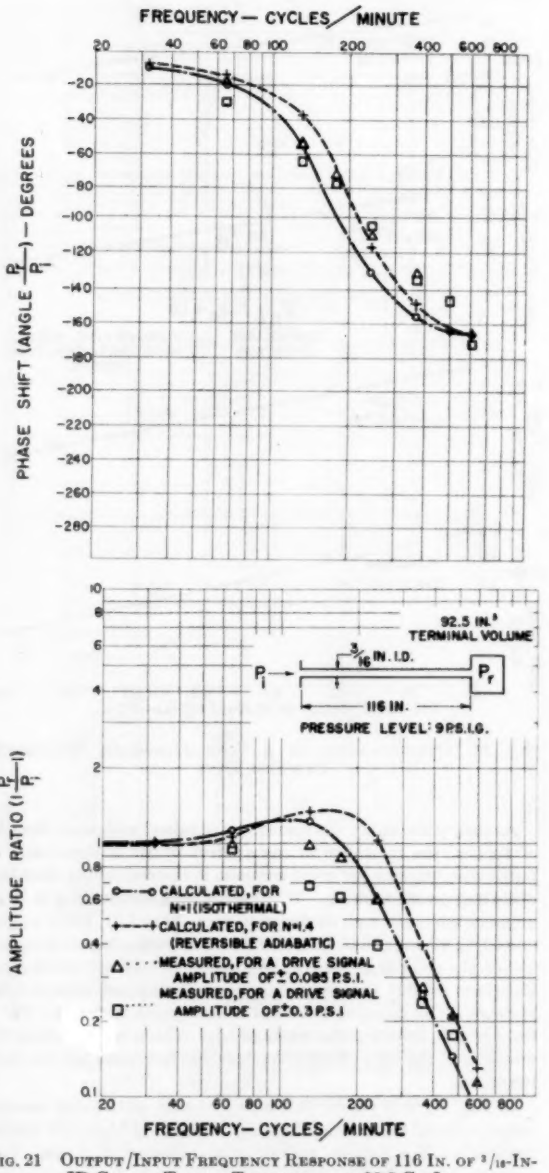


FIG. 21 OUTPUT/INPUT FREQUENCY RESPONSE OF 116 IN. OF $^3/_{16}$-IN-ID COPPER TUBING TERMINATED IN 92.5 CU IN.

as a limiting value with increasing frequency. Practically, for $^3/_{16}$-in-ID tubing at temperature and pressure levels of 59 F and 9 psig, respectively, α will remain essentially fixed at this value for all frequencies above 4 cps. Applying the values derived for the per unit length L-R-C parameters of standard $^3/_{16}$-in-ID tubing

$$\frac{R}{2}\sqrt{\frac{C}{L}} = 5.5 \times 10^{-4}\,\text{in.}^{-1}, \quad \text{for} \quad n = 1.4$$

Then, for tubing lengths exceeding 150 ft and at frequencies above 4 cps

$$\alpha l \gg 1; \quad e^{\alpha l} \gg e^{-\alpha l}$$

and Equation [17] reduces to $Z_i \approx Z_0$. Thus, for lengths of $^3/_{16}$-in-ID tubing exceeding 150 ft, the driving-point impedance will approach the characteristic impedance Z_0 of the tube with increasing frequency and hence tend to be independent of tubing length and the size of the terminating load. Theoretical calculations on standard $^3/_{16}$-in-ID tubing runs (dead ended or terminated in any size volume) indicate that the input loading effects, when the lengths exceed 200 ft, are essentially the same at frequencies above $^1/_2$ cps regardless of the termination volume

Further, noting that

$$Z_0 = \sqrt{\frac{L}{C}}\left(1 - j\frac{R}{\omega L}\right)^{1/2}$$

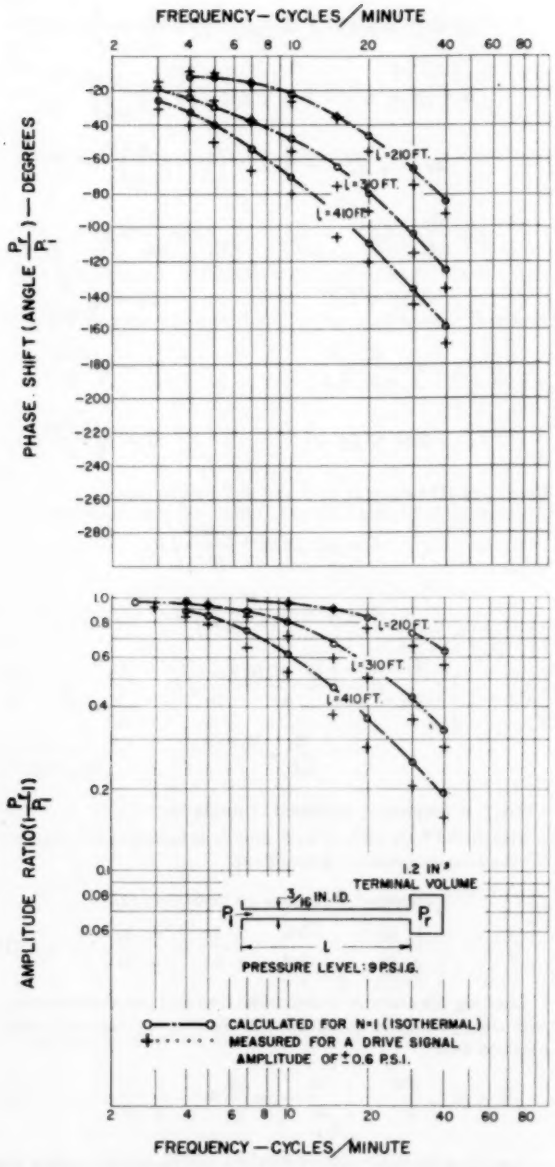


FIG. 22 OUTPUT/INPUT FREQUENCY RESPONSE OF 3/16-IN-ID DEAD-ENDED TRANSMISSION TUBING

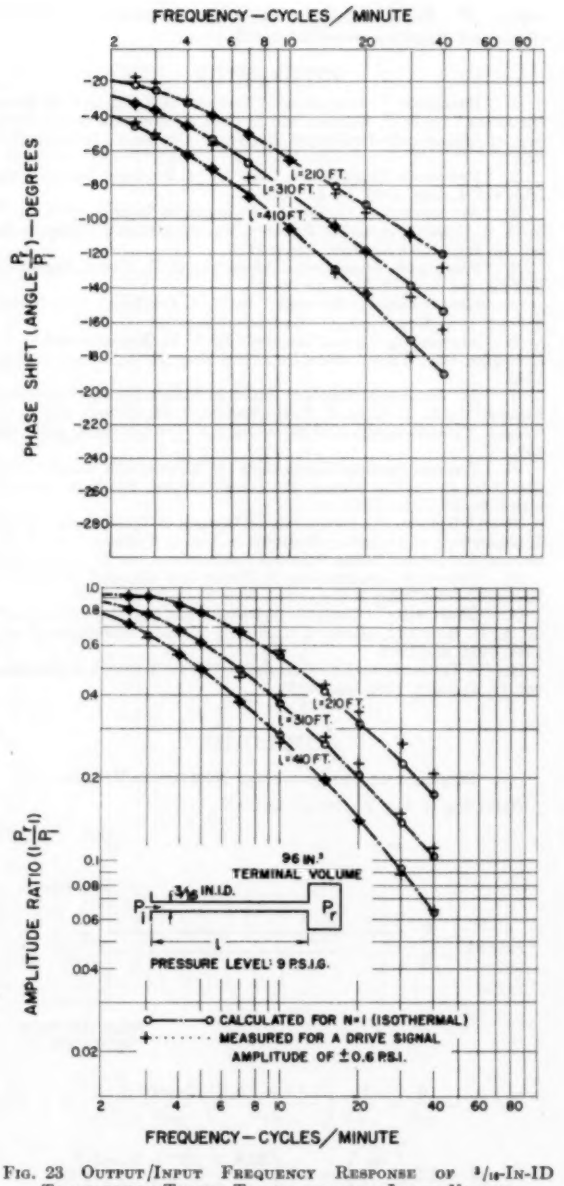


FIG. 23 OUTPUT/INPUT FREQUENCY RESPONSE OF 3/16-IN-ID TRANSMISSION TUBING TERMINATED IN A LARGE VOLUME

and

$$\lim_{\omega \to \infty} Z_0 = \sqrt{\frac{L}{C}} \quad \ldots\ldots\ldots\ldots\ldots\ldots [24]$$

the driving-point impedance of long lengths of tubing will approach the form of a resistance to "ground" (with a value of $\sqrt{L/C}$), with increasing frequency; while at low frequencies, the loading effect of tubing resembles that of a simple volume.

The Appendix presents a summary of the relations derived in this paper and gives numerical values for the parameters for 3/16-in-ID tubing at 59 F and 9 psig mean level.

ACKNOWLEDGMENTS

The authors wish to express their appreciation of the assistance given them on the test work and preparation of the paper by many members of the Research Department of the Brown Instruments Division of Minneapolis-Honeywell Regulator Company, especially by the following:

Dr. R. Shahbender, for his technical assistance during the preliminary phases of this investigation; Messrs. E. F. Hochschild and N. Yashin, for their assistance in collecting the test data; Messrs. W. A. Siegel, L. R. Leidy, R. D. Rea, and J. Izuka, for their assistance in performing the many calculations required for the establishment of the theoretical trends presented in this

paper; Mr. E. L. Kulick, for his skillful preparation of all the charts and diagrams presented in this paper.

BIBLIOGRAPHY

1 "Pneumatic Transmission of Instrument Readings Over Long Distances," by D. P. Eckman and L. Gess, Brown Instrument Division of Minneapolis-Honeywell Regulator Company, Bulletin No. B59-2.

2 "Pneumatic Transmission Lag," by M. Bradner, *Journal of the ISA*, vol. 4, July, 1949, pp. 618–625.

3 "Attenuation of Oscillatory Pressures in Instrument Lines," by A. S. Iberall, *Journal of Research*, National Bureau of Standards, vol. 45, July, 1950, Research Paper 2115.

4 "Pneumatic Transmission Lines," by J. C. Moise, *Journal of the ISA*, vol. 1, April, 1954, pp. 35–40.

5 "Communication Network," by E. A. Guillemin, John Wiley & Sons, Inc., New York, N. Y., first edition, vol. 11, 1935.

6 "Alternating Current Circuits," by R. M. Krechner and G. F. Corcoran, John Wiley & Sons, Inc., New York, N. Y., second edition, 1943.

7 "An Automatic Transfer Function Measuring and Recording System," by R. J. Ehret, E. F. Hochschild, J. M. Embrie, and E. C. Grogan, *Communication and Electronics*, November, 1953, pp. 664–669.

8 "Pressure Sensing Calculations for Aircraft and Guided Missiles," by J. D. Humphreys, Instrument Notes, Statham Laboratories, no. 25, June, 1953.

9 "A Note on the Evaluation of Designs of Transducers for the Measurement of Dynamic Pressures in Liquid Systems," by J. R. Barton, Instrument Notes, Statham Laboratories, no. 27, June, 1954.

10 "Electrical Transmission of Power and Signals," by E. W. Kimbark, John Wiley & Sons, Inc., New York, N. Y., 1949.

11 "Chart Atlas of Complex Hyperbolic and Circular Functions," by A. E. Kennelly, Harvard University Press, Cambridge, Mass., 1914, 1921, and 1924.

12 "A Transmission Line Calculator," by P. H. Smith, *Electronics*, vol. 12, January, 1939, pp. 27–32.

Appendix

Summary of Relations and Numerical Values

Referring to Fig. 24, for air at 59 F

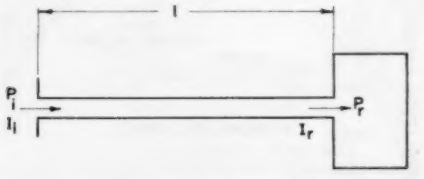


FIG. 24

$$L = \frac{\delta}{\pi r^2} = \frac{P_0}{14.7}\left(\frac{1.15 \times 10^{-7}}{\pi r^2}\right) \text{lb-sec}^2/\text{in.}^5 \Big/ \text{in.}$$

$$R = 1.2\left(\frac{8\mu}{\pi r^4}\right) = 1.2\left(\frac{20.8 \times 10^{-9}}{\pi r^4}\right) \text{lb-sec}/\text{in.}^5 \Big/ \text{in.}$$

$$C = \frac{\pi r^2}{nP_0} \text{ in.}^5/\text{lb} \Big/ \text{in.}$$

where

P_0 = mean pressure, psia
r = internal radius, in.
n = polytropic exponent

$$Z = R + j\omega L; \quad Y = j\omega C; \quad Y_r = \frac{1}{Z_r} = \frac{I_r}{P_r}$$

$$\alpha + j\beta = \sqrt{ZY} = j\sqrt{LC}\,\omega\left(1 - j\frac{R}{\omega L}\right)^{1/2}$$

$$Z_0 = \sqrt{Z/Y} = \sqrt{L/C}\left(1 - j\frac{R}{\omega L}\right)^{1/2}$$

For $\frac{R}{\omega L} < 0.6$

$$Z_0 \backsimeq \sqrt{L/C}$$

$$\alpha \approx \frac{R}{2}\sqrt{\frac{C}{L}}$$

$$\beta \approx \sqrt{LC}\,\omega$$

For $\frac{R}{\omega L} > 10$

$$Z_0 \approx \sqrt{\frac{R}{j\omega C}}$$

$$\alpha \approx \sqrt{\frac{RC\omega}{2}}$$

$$\beta \approx \sqrt{\frac{RC\omega}{2}}$$

For $^3/_{16}$-in-ID tubing at 59 F and 23.7 psia (9 psig)

$$L = 6.7 \times 10^{-6} \text{ lb-sec}^2/\text{in.}^5 \Big/ \text{in.}$$

$$R = 1 \times 10^{-4} \text{ lb-sec}/\text{in.}^5 \Big/ \text{in.}$$

$$C = \frac{1.15 \times 10^{-3}}{n} \text{ in.}^5/\text{lb} \Big/ \text{in.}$$

$$\frac{R}{\omega L} = \frac{2.4}{f}$$

where f = frequency, expressed in cycles/sec.

Figs. 5 and 6 are plots of α, β, and Z_0 as functions of frequency for the foregoing operating condition.

General Solution of Equation [1]

$$\frac{1}{\mu}\frac{dp}{dl} = -\frac{\partial^2 s}{\partial x^2} - \frac{1}{x}\frac{\partial s}{\partial x} + \frac{\delta}{\mu}\frac{\partial s}{\partial t} \quad \ldots\ldots\ldots [1]$$

Limiting the analysis to sinusoidal flow and pressure variations with time, $j\omega$ may be substituted for the $\partial/\partial t$ operator and vector notation used

$$\frac{\partial^2 S}{\partial x^2} + \frac{1}{x}\frac{\partial S}{\partial x} - j\omega\frac{\delta S}{\mu} = -\frac{1}{\mu}\frac{dP}{dl} \quad \ldots\ldots\ldots [25]$$

Assuming the fluid velocity S to be a continuous function of the distance x from the center of the tube

$$S = A_0 + A_1 x + A_2 x^2 + A_3 x^3 + \ldots \quad \ldots\ldots [26]$$

$$\frac{\partial S}{\partial x} = A_1 + 2A_2 x + 3A_3 x^2 + 4A_4 x^3 + \ldots \quad \ldots\ldots [27]$$

$$\frac{\partial^2 S}{\partial x^2} = 2A_2 + 6A_3 x + 12A_4 x^2 + 20A_5 x^3 + \ldots \quad [28]$$

$$\frac{P_r}{P_i} = \frac{2}{[e^{(\alpha+j\beta)l} + e^{-(\alpha+j\beta)l}] + Y_r Z_0 [e^{(\alpha+j\beta)l} - e^{-(\alpha+j\beta)l}]} \quad \ldots\ldots\ldots [16]$$

$$Z_i = \frac{P_i}{I_i} = Z_0 \left\{ \frac{[e^{(\alpha+j\beta)l} + e^{-(\alpha+j\beta)l}] + Y_r Z_0 [e^{(\alpha+j\beta)l} - e^{-(\alpha+j\beta)l}]}{Y_r Z_0 [e^{(\alpha+j\beta)l} + e^{-(\alpha+j\beta)l}] + [e^{(\alpha+j\beta)l} - e^{-(\alpha+j\beta)l}]} \right\} \quad \ldots [17]$$

Applying the equations to [25] and collecting terms

$$\frac{A_1}{x} + \left[4A_2 - j\omega\frac{\delta A_0}{\mu}\right] + \left[9A_3 - j\omega\frac{\delta A_1}{\mu}\right]x + \left[16A_4 - j\omega\frac{\delta A_2}{\mu}\right]x^2 + \left[25A_5 - j\omega\frac{\delta}{\mu}A_3\right]x^3 + \ldots = -\frac{1}{\mu}\frac{dP}{dl}$$

Equating terms on the left and right of the equation to evaluate the constants and substituting back into Equation [26]

$$S = j\frac{\omega\delta}{4\mu}A_0\left[\frac{4\mu}{j\omega\delta} + x^2 + \frac{j\omega\delta}{4\mu}\left(\frac{1}{2!}\right)^2 x^4 + \left(\frac{j\omega\delta}{4\mu}\right)^2\left(\frac{1}{3!}\right)^2 x^6 + \ldots\right] - \frac{1}{4\mu}\frac{dP}{dl}\left[x^2 + \frac{j\omega\delta}{4\mu}\left(\frac{1}{2!}\right)^2 x^4 + \left(\frac{j\omega\delta}{4\mu}\right)^2\left(\frac{1}{3!}\right)^2 x^6 + \ldots\right]$$

Applying the boundary condition that the fluid velocity is zero at the tube wall to evaluate A_0

$$S = \frac{dP}{dl}\left\{\frac{(r^2 - x^2) + \frac{j\omega\delta}{4\mu}\left(\frac{1}{2!}\right)^2(r^4 - x^4) + \frac{j\omega\delta}{4\mu}\left(\frac{1}{3!}\right)^2(r^6 - x^6) + \ldots}{4\mu + j\omega\delta r^2\left[1 + \left(\frac{j\omega\delta r^2}{4\mu}\right)\left(\frac{1}{2!}\right)^2 + \left(\frac{j\omega\delta r^2}{4\mu}\right)^2\left(\frac{1}{3!}\right)^2 + \ldots\right]}\right\} \quad \ldots\ldots [29]$$

The total volumetric flow rate I through a given section is found from

$$I = 2\pi\int_0^r Sx\,dx$$

$$I = 2\pi\frac{dP}{dl}\left\{\frac{\frac{1}{2}r^4\left[\frac{1}{2} + \frac{2}{3}\left(\frac{j\omega\delta r^2}{4\mu}\right)\left(\frac{1}{2!}\right)^2 + \frac{3}{4}\left(\frac{j\omega\delta r^2}{4\mu}\right)^2\left(\frac{1}{3!}\right)^2 + \ldots\right]}{4\mu + j\omega\delta r^2\left[1 + \left(\frac{j\omega\delta r^2}{4\mu}\right)\left(\frac{1}{2!}\right)^2 + \left(\frac{j\omega\delta r^2}{4\mu}\right)^2\left(\frac{1}{3!}\right)^2 + \ldots\right]}\right\}$$

Applying

$$R = \frac{8\mu}{\pi r^4};\quad X_L = j\omega\frac{\delta}{\pi r^2} = j\omega L$$

$$\frac{dP/dl}{I} = Z = R\left\{\frac{\frac{1}{2} + \frac{X_L}{R}\left[1 + \left(\frac{1}{2!}\right)^2\left(\frac{2X_L}{R}\right) + \left(\frac{1}{3!}\right)^2\left(\frac{2X_L}{R}\right)^2 + \left(\frac{1}{4!}\right)^2\left(\frac{2X_L}{R}\right)^3 + \ldots\right]}{\frac{1}{2} + \frac{2}{3}\left(\frac{1}{2!}\right)^2\left(\frac{2X_L}{R}\right) + \frac{3}{4}\left(\frac{1}{3!}\right)^2\left(\frac{2X_L}{R}\right)^2 + \frac{4}{5}\left(\frac{1}{4!}\right)^2\left(\frac{2X_L}{R}\right)^3 + \ldots}\right\}$$

Comparing the preceding equation with the relation

$$Z = R + \lambda X_L = R\left(1 + \lambda\frac{X_L}{R}\right)$$

and solving for λ

$$\lambda = \left[\frac{\frac{2}{3} + \frac{5}{6}\left(\frac{1}{2!}\right)^2\left(\frac{2X_L}{R}\right) + \frac{9}{10}\left(\frac{1}{3!}\right)^2\left(\frac{2X_L}{R}\right)^2 + \frac{14}{15}\left(\frac{1}{4!}\right)^2\left(\frac{2X_L}{R}\right)^3 + \ldots}{\frac{1}{2} + \frac{2}{3}\left(\frac{1}{2!}\right)^2\left(\frac{2X_L}{R}\right) + \frac{3}{4}\left(\frac{1}{3!}\right)^2\left(\frac{2X_L}{R}\right)^2 + \frac{4}{5}\left(\frac{1}{4!}\right)^2\left(\frac{2X_L}{R}\right)^3 + \ldots}\right] \quad \ldots\ldots [3]$$

As shown in Fig. 2, the foregoing may be approximated by

$$\lambda = 1 + \frac{1}{3}e^{-\left(\frac{\omega L}{4\pi R}\right)} \quad \ldots\ldots [4]$$

The important points in this approximation are that the phase angle of λ which is 5 deg or less may be treated as 0 deg for many purposes, and that λ varies in magnitude from 4/3 at $\omega \to 0$ to 1.0 at $\omega \to \infty$.

Discussion

R. P. Bigliano.[4] The writer wishes to compliment the authors on their paper which reduces this theory to a point where engineering solutions are practical.

The analytical functions provided and the plotted data will prove extremely useful to those process-control engineers in the chemical and related industries.

The authors exploit the electrical transmission-line theory to where it is possible to derive transfer functions and impedances from the physical-property data on the transmitting fluid and the dimensions of the system under investigation. This basic approach, along with the approximations made, have been shown to correlate with their experimental data on pneumatic systems.

The analysis can be extended and undoubtedly will find additional future applications as a process design tool. In process systems involving long pipelines fed by pumps and terminating in reactors it will be possible to design deliberately for minimum reflected energy by making the terminating impedance Z_R (reactor entrance design) equal to the characteristic impedance Z_0 of the process pipeline. In fact, it can be shown for the general transmission line that the reflection coefficient is

$$K = \frac{Z_R - Z_0}{Z_R + Z_0}$$

It is possible, from this single fact, to minimize or eliminate completely such things as water hammer in dense-liquid pipelines and pressure pulsations originating at a pump on the sending end of the line.

This paper provides an engineering tool which makes it possible to design a particular dynamic characteristic into the process. Industrial design and/or research departments should be encouraged to use and extend such analytical tools.

P. S. Buckley.[5] Although there have been several good papers on the theory of pneumatic transmission lines, this is the first which reduces the subject to useful, engineering proportions. The authors show both theoretically and experimentally that for small-amplitude sinusoidal driving signals one may predict the driving-point impedance and output-input transfer function with adequate accuracy by a simple, linear treatment which is almost identical mathematically to that used for electric transmission lines. Pneumatic-circuit elements—resistance, capacitance, and inertance—are all derived from readily available static data on tubing and load geometry, gas density, temperature, pressure, and viscosity. This means that much of electrical engineering technology of distributed systems may be transposed to the analysis and design of pneumatic transmission systems.

In the chemical and petroleum industry extensive use is being made of pneumatic control systems with an operating range of 3–15 psig the components of which are connected together by $^1/_4$-in-OD (sometimes $^3/_8$-in-OD) tubing. Process-control engineers in these industries will find considerable use for the plotted data on driving-point impedances and output-input transfer functions for various tubing lengths of 50 ft or greater with various loads.

For tubing lengths much less than 50 ft, the method of estimating resistance used in this paper leads to low values of resistance except for very small-amplitude driving signals. The reason is that for a given amplitude driving signal, the pressure drop per foot of tubing—and thereby the flow—is greater for a short length of tubing than for a long piece of tubing. Turbulent flow then results for smaller signals. The writer has found, therefore, that it is preferable to correlate resistance with the pressure drop per foot of tubing. In this way one may estimate the effective resistance for any signal amplitude and any line length.

This paper has interesting implications which go far beyond pneumatic transmission lines. The transmission of compressible fluids through pipelines is common in the chemical and petroleum industries. Hence, one may predict the dynamics of many process lines by the methods presented in this paper or by simple extensions thereof. The writer has found electric transmission-line theory of great help in analyzing distributed process systems.

A. R. Catheron.[6] The authors are certainly to be congratulated upon the thoroughness with which the study has been carried out, combining as it does a complete series of tests with a full theoretical analysis.

A few years ago, the writer had occasion to make tests quite similar in nature to those which form the foundation for the paper. The work was not carried beyond the test stage and no mathematical analysis was made. The data were plotted and filed for future use. When this paper appeared, the resemblance between many of the curves shown and those in our own test series was apparent, and results from both tests were replotted for better comparison in cases where the conditions were similar. All of our tests were made with an input amplitude of ± 0.1 psi, and the terminal volume (a pressure pickup) had a capacity of about 2 cc. Tubing sizes were the same, and the lengths were the same in the case of 100 ft (Fig. 12) and 50 ft (Fig. 14).

To refer to Fig. 12 of the paper, our tests show gain peaks at the same frequencies, with the curve running at, or slightly below, the test points given in the figure, the lower level being particularly apparent at the higher frequencies. The one exception appears at 130 cycles per min (cpm), where our gain reaches a level of 1.2, compared to 0.9 or 1.0. The phase curve also runs slightly below the points presented, a matter of 10 deg at the most. The Fig. 14 data were compared in the same way and, again, we confirm the authors' test data very closely, except for the peak at 235 cpm where our gain reaches 2.15 instead of 1.5 to 1.9. Our phase curve agrees with that of the paper with the exception of the change of slope between 350 and 600 cpm, which was much less pronounced. Our run was carried on to a frequency of 1000 cpm with a second gain peak (to 1.03) appearing at 850 cycles, and with a rapid increase in phase beyond 650 cycles reaching 320 deg at 960 cpm.

For comparison with Fig. 22, we had runs made with 200 ft, 300 ft, and 400 ft of line, respectively, again at ± 0.1 psi amplitude. Our results follow the pattern of the calculated and test values given in the paper, but with less attenuation in this case; as much as 30 per cent less at some frequencies. This seems reasonable since the tests for the paper were made at an amplitude of ± 0.6 psi. Our data cover a wider frequency range, to a gain of 0.06, which value occurred at frequencies of 125 cpm for 400 ft, 330 cpm for 300 ft, and 980 cpm for 200 ft. The phase curves show quite close coincidence with the calculated values of the paper, and they reach 320 deg at 95 cpm for 400 ft, 133 cpm for 300 ft, and 220 cpm for 200 ft. The corresponding gains are 0.135 at 95 cpm for 400 ft, 0.275 at 133 cpm for 300 ft, and 0.38 at 220 cpm for 200 ft.

Operation of process controls at frequencies above 1 cps has been rather uncommon in industrial practice—when it has occurred, there sometimes have been peculiarities in the performance. This demonstration that resonance and standing waves can exist in pneumatic lines should certainly help in clarifying these problems, and it will surely promote a better under-

[4] Research Engineer, Applied Physics, Engineering Department, Experimental Station, E. I. du Pont de Nemours & Company, Wilmington, Del.

[5] Research Associate, Chemical Engineering, E. I. du Pont de Nemours & Company, Wilmington, Del.

[6] The Foxboro Company, Foxboro, Mass.

standing of the principles of pneumatic transmission of data, as well as the unrealized possibilities of this technique.

F. D. Ezekiel.[7] The experimental results as well as the theoretical analysis that the authors present in the paper make a worth-while addition to the present knowledge of the dynamical behavior of pneumatic lines. In order to minimize the errors resulting from the introduction of lumped-system approximations, the writer wishes to submit the following technique for analyzing a pneumatic line. By considering such a line to be made up of a lossless, distributed-parameter conduit in series with a lumped-friction joint, the block diagram, Fig. 25, may be constructed.

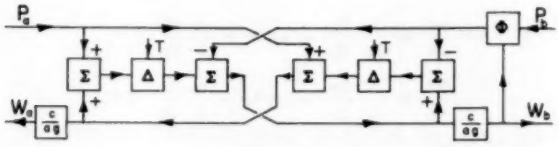


Fig. 25

This representation characterizes a hypothetical conduit in terms of a four-terminal network having two inputs P_a and P_b and two outputs W_a and W_b. The procedure by which this diagram has been obtained is outlined in detail in a paper by H. M. Paynter and the writer.[8]

The following nomenclature is used:

- a = line area, sq in.
- c = velocity of sound in air, ips
- g = acceleration of gravity, in/sec.^{2}
- P_a = upstream pressure, psia
- P_b = downstream pressure, psia
- W_a = upstream flow (into line, positive), lb/sec
- W_b = downstream flow (out of line, positive), lb/sec
- T = time delay equal to length of time it takes a pulse traveling at the speed of sound to get from one end of the line to the other, sec
- Δ = time-delaying component
- Σ = summing component
- Φ = component representing functional relationship between two variables

A. S. Iberall.[9] The reader of this paper can only assess the value of the treatment given with some knowledge of the historical background of the problem. Both the authors and the writer in his 1950 paper (authors' reference 3) on the same subject, have been remiss in supplying this background. To correct this the following sketches briefly some self-consistent historical highlights concerning the subject and the writer's comments on this paper.

1 Newton derived the velocity of propagation of an elastic wave as

$$V = \sqrt{\left(\frac{E}{\rho}\right)}$$

where

- V = velocity of propagation
- E = elastic modulus of medium
- ρ = density of medium

He applied this result to a gaseous medium to obtain

$$V = \sqrt{\left(\frac{p}{\rho}\right)} = C_N$$

where

- p = mean absolute pressure of medium
- C_N = this velocity has become known as the Newtonian velocity of sound

2 Laplace corrected this expression by pointing out that the waves of rarefaction and condensation will travel without "sensible loss in heat" (i.e., adiabatically) so that

$$V = \sqrt{\left(\frac{E_s}{\rho}\right)}$$

is the velocity of propagation where

- V = velocity of propagation
- E_s = adiabatic modulus

As applied to a gas, this results in

$$V = \sqrt{\left(\frac{\gamma p}{\rho}\right)} = C_0$$

where

- γ = ratio of specific heats
- C_0 = Laplacian velocity of propagation in an ideal gas

The mathematical implications of the velocity of propagation, in that it refers to propagation of a disturbance in a distributed medium, is that a wave motion in space is associated with a disturbance in time. Thus any measure of the disturbance (as for example, the pressure amplitude, or the particle velocity) is given by a function of the form

$$\psi = \psi(x \pm Vt)$$

in a one-dimensional problem; more generally

$$\psi = \psi(\bar{r} \pm \bar{V}t)$$

in a three-dimensional problem
where

- ψ = measure of a disturbed parameter (pressure, velocity, temperature, density, etc.)
- x = measure of distance in a one-dimensional problem
- $\bar{r}$ = vector measure of distance in a three-dimensional problem
- t = time

In particular problems we may consider functions of the form

$$\psi_i = A_i e^{a_i(x \pm Vt)}$$

or more generally, we may consider these as primitive generating functions from which solutions for specific boundary conditions may be built up. Thus the Laplacian result would require terms like

$$e^{a(x \pm C_0 t)}$$

where C_0 is used to denote $\sqrt{(E_s/\rho)}$.

In a problem involving a driving function (say the sinusoid $e^{j\omega t}$), then these functions take on the form

---

[7] Assistant Professor of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Mass. Assoc. Mem. ASME.

[8] "Computer Representations of Engineering Systems Involving Fluid Transients," by F. D. Ezekiel and H. M. Paynter, Paper No. 56—A-120.

"Effect of a Hydraulic Conduit with Distributed Parameters on Control Valve Stability," by F. D. Ezekiel, ScD thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Mass., 1955.

[9] Rand Corporation, Cleveland, Ohio. Mem. ASME.

$$e^{\frac{j\omega}{C_0}x \pm j\omega t}$$

We have thus assigned a value to a.

This result states that a sinusoidally driven wave is accompanied by a sinusoidal variation of "amplitude" in space. The quantity $j\omega/C_0$, or more generally, the factor that multiplies spatial distance, is called the "propagation constant." In this instance, it may be noted that the propagation constant is imaginary (i.e., real unattenuated space waves are produced).

This result is the basis of the elementary acoustic theory of plane waves, and thus, to some degree of success, can be applied to finite "plane waves" as one might encounter in a large diameter tube. Thus, for example, the nominal theory of organ pipes is accounted for.

3 However, it was pointed out by Stokes that even in a plane unbounded wave in a fluid, the equation

$$\frac{\partial^2\psi}{\partial t^2} = C_0^2\frac{\partial^2\psi}{\partial x^2}$$

is not obeyed, but, due to viscosity

$$\frac{\partial^2\psi}{\partial t^2} = C_0^2\frac{\partial^2\psi}{\partial x^2} + \frac{4}{3}\left(\frac{\mu}{\rho}\right)\frac{\partial^2\psi}{\partial x^2}$$

where

μ = viscosity of fluid medium

Solutions of this equation lead to terms of the form

$$e^{\pm\frac{j\omega}{C_0\sqrt{\left(1+\frac{4}{3}\frac{\mu}{\rho}\frac{\omega}{C_0^2}j\right)}}+j\omega t}$$

This result leads to a complex propagation constant

$$\Gamma = \frac{j\omega}{C_0\sqrt{\left(1+\frac{4}{3}\frac{\mu}{\rho}\frac{\omega}{C_0^2}j\right)}} \sim \frac{2}{3}\frac{\mu}{\rho}\frac{\omega^2}{C_0^2} + \frac{j\omega}{C_0^2}$$

for small ω or μ, Γ = propagation constant.

More generally, the definition of the propagation constant is written

$$\Gamma = \alpha + j\beta$$

in

$$e^{\Gamma x + j\omega t}$$

The quantity α is referred to as the "attenuation factor" (since it represents attenuation in space), and β is referred to as the "phase factor" (since it represents a phase shift). The quantity ω/β is referred to as the "phase velocity."

Thus Stokes' derivation showed that, even in free space, the possibility of an unattenuated plane acoustic wave no longer existed.

The elementary rudiments of transmission theory are well known and sketched out in books on electric transmission-line theory and acoustic-transmission theory. However, the basic problem, for application to pneumatic transmission lines, as was well pointed out by Taback in his NACA technical note, is to compute correctly the propagation constants of such a problem, since, different from the electrical case, they are not intrinsic properties of the line but of the line and fluid and of the hydrodynamic-flow regime.

4 Helmholz applied the ideas of the damping effect of viscosity to a wave limited in space (i.e., to one traveling in a tube), and then

5 Kirchoff proceeded to show that the effect of heat conductivity of the fluid is just as great as the effect of viscosity in modifying the propagation of waves in tubes. Since the more general result was obtained by Kirchoff, we may follow his results.

Assuming sufficiently small velocities, the complete set of linear equations that Kirchoff used were as follows:

(*a*) Mechanical equations of motion

$$\frac{\partial u}{\partial t} + \frac{1}{\rho_0}\frac{\partial p}{\partial x} = \nu_0\nabla^2 u + \frac{\nu_0}{3}\frac{\partial}{\partial x}(\nabla\bar{V})$$

$$\frac{\partial v}{\partial t} + \frac{1}{\rho_0}\frac{\partial p}{\partial y} = \nu_0\nabla^2 v + \frac{\nu_0}{3}\frac{\partial}{\partial y}(\nabla\bar{V})$$

$$\frac{\partial w}{\partial t} + \frac{1}{\rho_0}\frac{\partial p}{\partial z} = \nu_0\nabla^2 w + \frac{\nu_0}{3}\frac{\partial}{\partial z}(\nabla\bar{V})$$

(*b*) Equation of continuity

$$\frac{\partial s}{\partial t} = -\nabla\bar{V} = -\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} - \frac{\partial w}{\partial z}$$

(*c*) Equation of energy

$$\frac{\partial T}{\partial t} = (\gamma - 1)T_0\frac{\partial s}{\partial t} + \frac{k}{\rho_0 C_v}\nabla^2 T$$

(*d*) Equation of state for an ideal gas

$$\frac{p}{p_0} = s + \frac{T}{T_0}$$

where

u, v, w = components of velocity
s = condensation [$\rho = \rho_0(1 + s)$]
T = absolute temperature
ν_0 = kinematic viscosity ($= \mu_0/\rho_0$)
k = thermal conductivity
c_v = specific heat at constant volume

From these equations, he derived what has become known as the Kirchoff equations of sound

$$h^2\theta - \left[C_0^2 + h\left(n + \frac{4}{3}\nu_0\right)\right]\nabla^2\theta + \frac{n}{h}\left[C_N^2 + \frac{4}{3}h\nu_0\right]\nabla^4\theta = 0$$

Here

$$\theta = \frac{T}{(\gamma - 1)T_0}$$

$$h = j\omega$$

$$n = \frac{k}{\rho_0 C_v}\ (= \text{thermometric conductivity})$$

Kirchoff applied this equation to driven waves contained in a tube under the assumption of small viscosity to obtain

$$\Gamma = \frac{j\omega}{C_0}\sqrt{\left\{1 + \frac{2\left[\sqrt{\nu_0} + \left(\gamma - \frac{1}{\gamma}\right)\sqrt{n}\right]}{r\sqrt{h}}\right\}}$$

where r = radius of tube.

6 On the other hand Rayleigh applied the Kirchoff equations to the case of large viscosity to obtain

$$\Gamma = \sqrt{\left(\frac{8\nu_0 w}{C_N{}^2 r^2}\, j\right)}$$

It was soon recognized that the sharp discriminant between these two cases (small viscosity versus large viscosity, or large tube versus small tube; or high frequency versus low frequency) was the quantity

$$r\sqrt{\left(\frac{\omega}{\nu_0}\right)}$$

In the writer's 1950 paper, the following definition was used

$$z = \frac{r^2\omega}{\nu_0}$$

If z is less than 1 and the velocity low enough, the tube is narrow, the Rayleigh result holds, the inertia may be negected, Poiseuille's law for the velocity distribution holds, there is laminar and lamellar flow in the tube, the system is overdamped and the flow is isothermal!

It is of interest to quote Rayleigh's discussion of this case:[10]

"One result of the investigation may be foreseen. When the diameter of the tube is very much reduced, the conduction of heat from the center to the circumference of the column of air becomes more and more free. In the limit the temperature of the solid walls controls that of the included gas, and the expansions and rarefactions take place isothermally. Under these circumstances there is no dissipation due to conduction, and everything is the same as if no heat were developed at all. Consequently the coefficient of heat-conduction will not appear in the result, which will involve, moreover, the Newtonian value of the velocity of sound, and not that of Laplace."

The Rayleigh result was to be expected in the sense that for slow enough oscillations, the static result (Poiseuille's law) should hold at each point in the tube. The only question at issue would be at what value of the parameter of the system would this regime break down. This can and has been answered in a number of ways, all equivalent to the statement that z must be less than 1.

On the other hand the Kirchoff result did not prove out so easily by experimental investigation. The basic experimental requirement was that the result of the Kundt-tube experiment (sonic resonance in a tube) should be correctable to obtain the Laplacian free-air velocity of sound. Since from a strictly theoretical point of view, the validity of the sciences of the mechanics of continuous media, of hydrodynamics, and more particularly of statistical mechanics hung in the balance, it was necessary that the Kirchoff results be given searching inquiry. It was so tested (many references discuss it exhaustively.)[11] However, by 1933 the major elements of controversy over the Kirchoff result were fairly well resolved. The following is quoted (slightly out of context, avoiding some minor arguments) from one of the latest and most sound investigations.[12] In summary the authors state:

"The results indicate that for all the gases (tested), the Helmholtz-Kirchoff formula is quantitatively correct in its statement of the influence of tube diameter and frequency on the velocity"; and "An examination of the literature indicates that much of the criticism to which the Helmholz-Kirchoff formula has been subjected in the past was unjustified"; and "the following free space velocities for the pure dry gases have been deduced (by the H-K equations) from the observations. The probable accuracy is of the order of 1 in 1000 except for hydrogen where it may be a little less."

Thus the Kirchoff results were known to be correct for large tubes (for values of z greater than 100) and the Rayleigh results were known to be correct for small tubes (for values of z less than 1).

What then remained to be done? Somehow in transition as a function of

$$z = \frac{r^2\omega}{\nu_0}$$

a solution was required that went over continuously between these extremes.

7 Illustrative of common acoustic solutions is one taken from Crandall.[13] The equation of motion is taken as

$$\left[j\omega\rho_0 - \frac{\mu_0}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right)\right]u = -\frac{\partial p}{\partial x}$$

(The same equation as assumed by Rohmann and Grogan) with the solution

$$u = \frac{1}{j\omega\rho_0}\frac{\partial p}{\partial x} + AJ_0(Kr)$$

$$K^2 = -j\frac{\omega}{\nu_0}$$

where

J_0 = Bessel function of zero order
A = an undetermined constant

Applying the boundary condition of vanishing velocity at the wall

$$u(r) = -\frac{1}{j\omega\rho_0}\frac{\partial p}{\partial x}\left[1 - \frac{J_0(Kr)}{J_0(Kr_0)}\right]$$

where

r_0 = tube diameter
r = here is the radial variable

The mean velocity (over the section) becomes

$$\bar{u} = -\frac{1}{j\omega\rho_0}\frac{\partial p}{\partial x}\left[1 - \frac{2J_1(Kr_0)}{Kr_0J_0(Kr_0)}\right]$$

(The identical result that is given in the writer's paper, authors' reference 3, as Equation [75].)

If the impedance is defined by

$$\bar{u} = -\frac{1}{Z}\frac{\partial p}{\partial z}$$

(see for example the text by R. W. P. King[14]), then

$$Z = \frac{j\omega\rho_0}{1 - \dfrac{2J_1(Kr_0)}{Kr_0J_0(Kr_0)}}$$

[10] "Theory of Sound," by Lord Rayleigh, The Macmillan Company, London, England, second edition, 1894–1896, sec. 350.

[11] See for example: "The Specific Heat of Gases," by J. R. Partington and W. G. Schilling, D. Von Nostrand Company, Inc., New York, N. Y., 1924.

[12] "The Velocity of Sound in Gases in Tubes," by G. C. W. Kaye and G. G. Sherratt, Proceedings of the Royal Society of London, England, series A, vol. 141, 1933, pp. 123–143.

[13] "Theory of Vibrating Systems and Sound," by I. B. Crandall, D. Van Nostrand Company, Inc., New York, N. Y., 1927.

[14] "Transmission Line Theory," by R. W. P. King, McGraw-Hill Company, Inc., New York, N. Y., 1955, p. 48.

(A minor note should be taken of two means of choosing the normalization for impedance. One choice of impedance is based on the line, another on the medium. Thus different authors' results may differ by a factor of the area of the tube πr_0^2.)

For small values of the argument $|Kr_0|$

$$J_0(x) = 1 - \frac{x^2}{4} + \frac{x^4}{64}$$

$$J_1(x) = \frac{x}{2}\left(1 - \frac{x^2}{8} + \frac{x^4}{192}\right)$$

so that

$$Z = \frac{8\mu_0}{r_0^2} + \frac{4}{3} j\omega\rho_0$$

$$= R + jX$$

where

$$R = \frac{8\mu_0}{r_0^2}$$

$$X = \frac{4}{3}\omega\rho_0$$

This result agrees with the Rayleigh result.

For large values of the argument $|Kr_0|$

$$\frac{2J_1(x)}{xJ_0(x)} = \frac{1}{x^2} - \frac{2j}{x}$$

so that

$$Z = j\omega\rho_0\left[1 - \frac{2j}{Kr_0}\right]$$

This agrees with the Kirchoff result.

Unfortunately, the approximate methods and assumptions that Crandall gives run into difficulty in assuring a correct value for the propagation constant.

8 Quite probably the most rigorous completion of Kirchoff's investigation was given in the writer's 1950 paper. The details of the theoretical program undertaken cannot be given briefly. Suffice it to say that the equation set chosen was complete; that is, it included the complete three-dimensional linearized (small-amplitude) equation of motion, continuity, energy, and the assumption of the existence of an equation of state. The rigorous treatment of this investigation established another self-consistent criterion for the validity of the results; namely, that the flow regime pertains to continuous media in which equipartition of energy is assured. Mathematically these criteria were that

$$\frac{\nu_0}{2C_0 r_0} < 1$$

$$\frac{\nu_0\omega}{C_0^2} < 1$$

With these two added restrictions, a result was obtained which went over continuously from the Rayleigh to the Kirchoff results.

The "exact" equations given by the writer are contained in Equations [74] and [75] in his paper; namely

$$Z = \frac{j\omega\rho_0}{1 - \dfrac{2J_1(Kr_0)}{Kr_0 J_0(Kr_0)}}$$

$$\Gamma^2 = -\frac{\omega^2}{C_0^2}\frac{1 + (\gamma - 1)\dfrac{2J_1(lr_0)}{lr_0 J_0(lr_0)}}{1 - \dfrac{2J_1(Kr_0)}{Kr_0 J_0(Kr_0)}}$$

where

$$K^2 = -j\frac{\omega}{\nu_0}$$

$$l^2 = -j\frac{\omega}{\nu_0}\sigma_0$$

where

$$\sigma_0 = \text{mean Prandtl number}\left(= \frac{c_p\mu_0}{k}\right)$$

To the best of the writer's knowledge, the first rigorous expression for the attenuation parameter was derived for his 1950 paper.

The explicit statement of the restrictions under which these results hold are as follows:

(*a*) Small maximum Reynolds number (below 2000).

(*b*) Any elastic fluid in which

$$\frac{\nu_0}{2C_0 r_0} < 1$$

and driven at a frequency such that

$$\frac{\nu_0\omega}{C_0^2} < 1$$

(*c*) Long tube (i.e., above a certain minimum length his results will hold. However, end corrections may or may not be significant).

Under these restrictions, the writer's paper gives the "exact" answer. Investigators may either attempt by elementary means to derive these results, may use his results as the basic for appropriate approximations, or more significantly derive results that obtain (and these are needed) where the restrictions he has stated are violated.

Since the publication of the paper in 1950, the writer has taken exception to all efforts that in his opinion are a step backward on this subject. The results of Rayleigh and Kirchoff are no longer to be questioned, and the rational transition between these results given in the writer's paper are also no longer to be questioned theoretically (they are only open now to experimental challenge).

Thus on the basis of this historical introduction one may criticize the present paper.

The authors' equation of motion

$$\left[j\omega\rho_0 - \frac{\mu_0}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right)\right]u = -\frac{\partial p}{\partial x}$$

is given in standard books in acoustics. By virtue of a fortunate set of compensations it can give correctly the impedance of a long line to be

$$Z = \frac{j\omega\rho_0}{1 - \dfrac{2J_1(Kr_0)}{Kr_0 J_0(Kr_0)}}$$

Instead of integrating the equation correctly by means of Bessel functions, the authors have chosen an approximate integration which leads to

$$Z = \frac{8\mu_0}{r_0^2} + \left[1 + \frac{1}{3}e^{\frac{j(Kr_0)^2}{32\pi}}\right]j\omega\rho_0$$

For small values of the argument Kr_0 both these expressions give

$$Z = \frac{8\mu_0}{r_0^2} + \frac{4}{3}\omega\rho_0 j$$

(i.e., the exponential term is unnecessary).

For large values of the argument the rigorous expression leads to

$$Z = j\omega\rho_0\left[1 - \frac{2j}{Kr_0}\right]$$

whereas the authors' result leads to

$$Z = \frac{8\mu_0}{r_0^2} + \frac{1}{3}\, j\omega\rho_0 e^{\frac{j(Kr_0)^2}{32\pi}}$$

$$= \frac{8\mu_0}{r_0^2} + \frac{1}{3}\, j\omega\rho_0 e^{\frac{\omega\rho_0 r_0^2}{32\mu_0}}$$

which is nonsense.

Thus their basic argument can only be that they have chosen an (unjustified) approximation which permits some indeterminate expansion beyond $z = 1$. The choice of the elegance of an exact result permitting simple tabular solution, as compared to a one or two-term power expansion is not to be resolved simply in favor of the latter. (A comparable argument that the tedium of using tables of sin x makes power solutions of the form $x - x^3/3$ more attractive is certainly ridiculous for large values of the argument, and only warranted when attention is always focused on small values of the argument.)

When it comes to the propagation constant, their arguments are meaningless. The authors disregarded the warning that Rayleigh gave and that the writer gave in his 1950 paper that their program is only valid if the Poiseuille regime is assumed, and in that case the flow is isothermal. They have disregarded the thought that if the flow is to be assumed "polytropic" the coefficient of the polytropic expansion may have to be taken complex and not simply a real number between 1 and γ. Thus the only real recourse they have in their result is to compare it with the writer's.

The rigorous result for the propagation constant is

$$\Gamma^2 = -\frac{\omega^2}{C_0^2}\,\frac{1 + (\gamma - 1)\,\dfrac{2J_1(lr_0)}{lr_0J_0(lr_0)}}{1 - \dfrac{2J_1(Kr_0)}{Kr_0J_0(Kr_0)}}$$

where

$$K^2 = -j\,\frac{\omega}{\nu_0}$$

$$l^2 = -j\,\frac{\omega}{\nu_0}\,\sigma_0$$

In the 1950 paper the writer stated that a goodly approximation to this result can be taken by assuming that either the Prandtl number is sufficiently close to 1 (for gases) or that γ is sufficiently close to 1 (for liquids), that one may write

$$\Gamma^2 = -\frac{\omega^2}{C_0^2}\,\frac{1 + (\gamma - 1)\,\dfrac{2J_1(Kr_0)}{Kr_0J_0(Kr_0)}}{1 - \dfrac{2J_1(Kr_0)}{Kr_0J_0(Kr_0)}}$$

This similarity makes the propagation constant a function only of the same argument Kr_0 as is the impedance.

The authors' result is

$$\Gamma^2 = -\frac{\omega^2}{mC_N^2}\left[1 - 1.2\,\frac{8}{(Kr_0)^2}\right]$$

For small values of the argument the "exact" result is

$$\Gamma^2 = \frac{\omega^2}{C_0^2}\,\frac{8\gamma}{(Kr_0)^2}$$

The authors' result is

$$\Gamma^2 = \frac{1.2\omega^2}{mC_N^2}\,\frac{8}{(Kr_0)^2}$$

so that it would be necessary to assign the value $m = 1.2$, instead of the unit value they assigned. The basically correct answer is that for their assumptions the coefficient that was given the value 1.2 should be 1 and the polytropic coefficient m should be 1.

For large values of the argument the exact result is

$$\Gamma^2 = -\frac{\omega^2}{C_0^2}\left[1 - \frac{2j\gamma}{Kr_0}\right]$$

instead of

$$\Gamma^2 = -\frac{\omega^2}{mC_N^2}\left[1 - 1.2\,\frac{8}{(Kr_0)^2}\right]$$

Here the authors would have to assume the value $m = \gamma$. However, it should be noted that their deviation from unity is not of the correct order.

In summary, thus lending the authority of publication in an ASME Journal to such poorly drawn theory is in the writer's opinion an unwarranted burden on the engineering profession, particularly in the light of the relative completeness of previous developments. It is to be regretted that the authors and reviewers have not seen fit to acquaint themselves adequately with the literature.

Authors' Closure

Mr. Bigliano and Mr. Buckley have suggested interesting extensions to the use of the technique of this paper. We are in agreement with Mr. Buckley's method of examining pressure drop per foot of tubing when estimating the likelihood of turbulence.

Mr. Catheron's comments and addition to the data are greatly appreciated and add to the value of the total material presented.

Dr. Ezekiel has presented a method for computer simulation of pneumatic lines that is especially valuable in systems where low-order lumped approximations do not give the desired accuracy. The higher-order approximations derived as described in the paper under discussion become very awkward to handle on an analog computer.

Mr. Iberall is to be thanked for his extensive addition to the bibliography and background material. However, he seems to have missed the entire point of the paper. The purpose has been to present a relatively simple approximate analysis leading to a solution in a convenient form for engineering computations. At the same time good correlation with test data was required. In his discussion, Mr. Buckley has pointed out a principal area of application of this paper. The particular form chosen makes the theory and familiarity built up for electric transmission lines applicable to pneumatic transmission. It also makes the treatment of the transmission lines as one interacting component of a dynamic system practical.

In this paper the equation of motion has been solved to yield the form $dp/dl = I(R + \lambda Y_L)$ with all symbols as defined in the nomenclature. The dimensionless vector quantity λ is plotted in Fig. 2. Since λ has both varying magnitude and phase, both the resistance and the inertance are varying quantities if they are defined as the in-phase and quadrature components, respectively. As explained in the paper, it has been found sufficient for engi-

neering purposes to treat both these quantities as constants and hence simplify their use. The error introduced by this procedure is discussed in the body of the paper. The method of development has been chosen to make it easy to see the order of approximation being made by this procedure. Note that the quantity $\omega L/R$ used in discussing the parameters and limit cases in the paper is $\omega r_0^2/\nu_0$ in Mr. Iberall's notation.

Since it was decided to hold R constant with frequency when it is known to vary slightly, the correction factor 1.2 was applied to the theoretical minimum resistance so as to give the best correlation between test data and the working equations over a large number of tests. With the method and purpose of this paper in mind, the correlation between the two methods discussed by Mr. Iberall will now be considered briefly. First the quantity Z Mr. Iberall discusses is not the impedance of a long line but the impedance per unit length where the compressibility effects can be neglected. If the solution of this paper is written in Mr. Iberall's notation, the solution is

$$Z = \frac{8\mu_0}{r_0^2} + \left[1 + \frac{1}{3} e^{\frac{-j(Kr_0)^2}{32\pi}}\right] j\omega\rho_0$$

with the sign changed from that mistakenly given by Mr. Iberall. This leads to the limit cases

$$Z = 8\mu_0/r_0^2 + (4/3)j\omega\rho_0 \qquad \text{for } (Kr_0) \text{ small}$$

and

$$Z = 8\mu_0/r_0^2 + j\omega\rho_0 \qquad \text{for } (Kr_0) \text{ large}$$

As noted in the paper, it has been found satisfactory to neglect the 4/3 factor and treat it as unity since the resistance term predominates in the region where the 4/3 factor is applicable. It can be seen that for both the impedance term and the propagation constant squared, which Mr. Iberall chooses to discuss, the predominant terms are in agreement if the polytropic exponent is correctly taken. Of course, the empirical factor used to center the resistance parameter in the working range cannot be expected to check either limit case. It is the authors' opinion that separating the real and imaginary components of the propagation constant for the limit cases provides better insight into the behavior of pneumatic lines. This has been done in the Appendix.

It is correct that for the limit case at low frequency with a full Poiseuille velocity distribution, the action is isothermal. This should not be extended mistakenly to mean that whenever the flow is laminar it is isothermal. The usual definition of a "polytropic" expansion is one for which the equation is

$$pv^n = K, \text{ a real constant}$$

where p and v are the pressure and specific volume, respectively, state functions of the fluid. For the suggestion of a complex exponent n to have meaning, a new definition would be required. It should be noted that in usual engineering cases the existence of a polytropic action is empirically determined. Only the special cases of isothermal and reversible adiabatic action can be rigorously derived. Figs. 17 and 18 provide data against which to check the derived propagation constants.

In connection with the oral presentation, the authors demonstrated one application of this paper. On numerous other occasions this method has been used in system analysis to good advantage. It is the good correlation between test and calculated results, together with the ease of handling inherent in the form of solution chosen, that gives this method value. On the basis of contact with the other reviewers, as well as the method under discussion, the authors are forced to conclude that Mr. Iberall's closing remarks are inane.

The authors wish to thank the discussers for their interest and suggestions and the Society for the considerable space devoted to the publication of this paper.

Analysis and Design of a Servomotor Operating on High-Pressure Compressed Gas[1]

By GERHARD REETHOF,[2] DETROIT, MICH.

The analysis, design, and development of a high-performance gas servomechanism suitable for aircraft and missile applications are described. The flow-control valve consists of two upstream orifices of fixed areas and two differentially variable downstream orifices. The areas of the variable orifices are controlled by a simple electromagnetic actuator. The fluid motor consists of two single-acting self-lubricating pistons which actuate the rocker-arm load assembly through push rods. Design parameters were obtained from the results of an analog-computer design study. The predicted transient response agrees very well with the experimentally derived rise time of 4 millisec.

Nomenclature

The following nomenclature is used in the paper:

A = area, sq in.
A_D = flow area of variable downstream orifice, sq in.
A_{D0} = downstream-orifice mid-position flow area, sq in.
A_f = area of approach passage to downstream orifice, sq in.
A_n = flapper pressure-sensitive area, sq in
A_R = ram area, sq in.
A_u = upstream-orifice area, sq in.
b_L = load-damping coefficient, lb-sec/in.
C_1 = weight-flow coefficient, $\sqrt{F}$/sec
D = differential operator d/dt, 1/sec
D_f = diameter of approach passage to downstream orifice, in.
dm = differential mass of gas, lb-sec^2/in.
e = error-signal voltage, volt
e^* = command signal, volt
e_L = voltage proportional to load position, volt
F_m = force developed by valve actuator, lb
F_L = load disturbance, lb
g = gravitational acceleration, in/sec^2
K = load spring, lb/in.
K_a = gain of d-c amplifier, amp/volt
K_f = position-feedback gain, volt/in.
K_v = velocity-feedback gain, volt-sec/in.
K_1 = current sensitivity of valve actuator, lb/amp
K_3 = effective stiffness of valve actuator, lb/in.
k = ratio of specific heats
m_L = load mass, lb-sec^2/in.
P_L = load pressure $P_1 - P_2$, psi
P_s = supply-source pressure, psi
P_0 = zero-load cylinder pressure with valve in mid-position, psi
P_1 = pressure on ram 1, psi
P_2 = pressure on ram 2, psi
$\mathcal{P}$ = power, lb-in/sec
$\mathcal{P}_0$ = nondimensionalizing power, lb-in/sec
Q_m = volume flow to motor, cu in/sec
Q_{m0} = nondimensionalizing volume flow to motor, cu in/sec
q_m = nondimensional volume flow to motor
R = gas constant, sq in/sec^2 deg F abs
T = temperature, deg F abs
T_s = supply-gas temperature, deg F abs
T_1, T_2 = gas temperatures in chambers 1 and 2, deg F abs
u = specific internal energy of gas, in.
V_0 = volume under compression at motor mid-position, cu in.
V_1, V_2 = volumes of motor chambers 1 and 2, cu in.
v = specific volume of gas, cu in/lb
W = weight flow of gas, lb/sec
W_f = flow through downstream orifices, lb/sec
W_m = motor flow, lb/sec
W_s = supply flow, lb/sec
x = valve displacement from mid-position, in.
x_0 = mid-position opening of valve, in.
z = load position, in.
$\dot{z}$ = load velocity, in/sec
$\ddot{z}$ = load acceleration, in/sec^2
α = orifice-area ratio, A_{D0}/A_u
γ = nondimensionalized valve displacement x/x_0
Δ = incremental change in quantity
ζ = damping ratio
ρ = weight density of gas, lb/cu in.
ω_n = undamped natural frequency of open-loop servomotor, radians/sec

Introduction

The utilization of compressed air for the actuation of control valves in the field of process control is widespread and well known. The air-pressure supply used to actuate these controls seldom exceeds 150 psig. The relatively low speed of response of these control systems has led to the belief that pneumatic controls could not be designed to meet the much higher speed-of-response requirements of aircraft, missile, and machine-tool applications. Until recently, the highly refined oil-hydraulic control systems have been used to meet the requirements of these applications. Owing to the constantly increasing temperature requirements for missile and aircraft components and the high operating temperatures of nuclear-reactor controls, interest in the development of high-speed pneumatic control devices has been renewed within recent years.

In this paper the basic analysis and subsequent design of a

[1] The work reported was performed at the Dynamic Analysis and Control Laboratory, Massachusetts Institute of Technology, and supported by the United States Air Force under Contract No. AF 33-(616)-2356. This paper is based in part on an ScD thesis, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Mass., 1955.

[2] Vickers, Inc.; formerly, Assistant Professor of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Mass. Mem. ASME.

Contributed by the Instruments and Regulators Division and presented at a joint session with the Machine Design Division at the Semi-Annual Meeting, Cleveland, Ohio, June 17–21, 1956, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, January 26, 1956. Paper No. 56—SA-20.

specific pneumatic servomotor are presented. The purpose of the treatment is to demonstrate a successful approach to the design of high-speed pneumatic control systems.

Description of the System

The system is illustrated schematically in Fig. 1. The flow valve is supplied with a high-pressure gas from a constant-pressure gas-power source. The power-control valve consists of two upstream orifices of fixed area and two downstream orifices with differentially variable areas. The control pressures in the chambers between the upstream and downstream orifices are piped to each side of the fluid motor. In this instance, the fluid motor consists of two single-acting, self-lubricating pistons which, through two spherical-ended connecting rods, apply a torque to the rocker arm. The rocker arm can be connected to the load by means of a splined shaft. The position of the output member is measured by a linear differential transformer which delivers an a-c signal. This signal is rectified in a keyed demodulator and fed back to the adding circuit of the forward-loop d-c amplifier. The push-pull output stage of the amplifier supplies the polarizing currents and differentially varying signal currents to the two coils of the E-style electromagnetic valve actuator. The small motion of the plugs on the rotor of the valve actuator control the flow and pressure conditions in the ram-type motor. An I-beam flexure-pivot mounting between rotor and stator ensures frictionless motion of the rotor and maintains correct spacing of the air gaps. Since the magnetic fields produce a statically unstable valve actuator, a leaf spring is mounted on the stator and connected to the rotor by wire links. A constant tension in the wire links prevents backlash in the motion of the rotor.

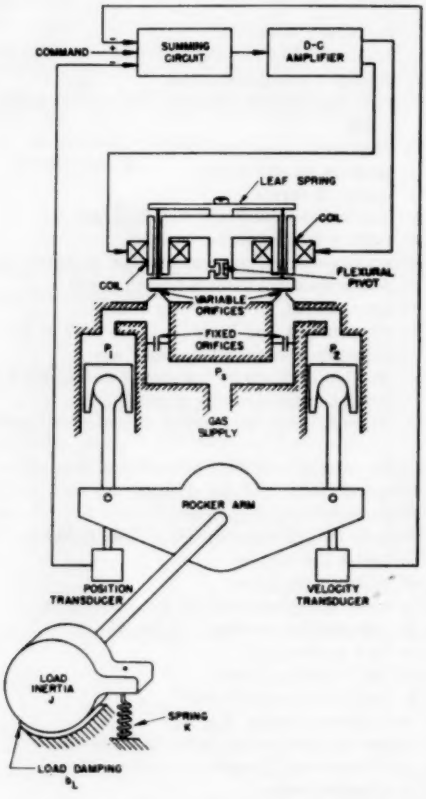


Fig. 1 Schematic Diagram of System

The desired high speed of response of the complete control system is obtained by feeding back a signal proportional to the velocity of the output member through an amplifier with an adjustable gain.

Component Analysis

Flow Valve. The control of the flow of fluid power from the gas generator to the motor-load system is accomplished by a four-way valve consisting of two three-way valves operating in a push-pull arrangement. The analysis of the weight flow through the orifices is based on the assumption of frictionless adiabatic flow of a perfect gas (1).[1]

For a given upstream pressure and temperature, the ratio of weight flow per unit area to the maximum weight flow per unit area is

$$\frac{\frac{W}{A}}{\left(\frac{W}{A}\right)_{max}} = \sqrt{\frac{2}{k-1}\left(\frac{k+1}{2}\right)^{\frac{k+1}{k-1}}\left(\frac{P}{P_s}\right)^{\frac{1}{k}}}\sqrt{\left(1-\frac{P}{P_s}\right)^{\frac{k-1}{k}}} \quad \ldots\ldots [1]$$

This relationship will be denoted by A_{max}/A.

For orifice flow with pressure ratios below the critical pressure ratio, given by Equation [2], the value of A_{max}/A remains at 1. Therefore, for the condition of critical flow ($A_{max}/A = 1$), the weight flow is independent of downstream conditions

$$\frac{P}{P_s} = \left(\frac{2}{k+1}\right)^{\frac{k}{k-1}} \quad \ldots\ldots [2]$$

For a specific gas, the weight-flow equation becomes

$$W = C_1 A \frac{P_s}{\sqrt{T_s}}\left(\frac{A_{max}}{A}\right) \quad \ldots\ldots [3]$$

where

$$C_1 = g\sqrt{\frac{k}{R\left(\frac{k+1}{2}\right)^{\frac{k+1}{k-1}}}} \quad \ldots\ldots [4]$$

Three-Way Valve. From geometric considerations, the downstream-orifice area is given by

$$A_D = \pi D_f(x_0 - x) \quad \ldots\ldots [5]$$

With the valve at mid-position ($x = 0$), the orifice area is

$$A_{D0} = \pi D_f x_0 \quad \ldots\ldots [6]$$

Division of Equation [5] by Equation [6] results in the dimensionless equation

$$\frac{A_D}{A_{D0}} = 1 - \frac{x}{x_0} \quad \ldots\ldots [7]$$

The weight flow through the downstream variable orifice for the expected critical-flow conditions is given by

$$W_{f1} = C_1 A_{D0}\left(1 - \frac{x}{x_0}\right)\frac{P_1}{\sqrt{T_s}} \quad \ldots\ldots [8]$$

The weight flow through the upstream fixed-area orifices is given by

[1] Numbers in parentheses refer to the Bibliography at the end of the paper.

$$W_{s1} = C_1 A_u \frac{P_s}{\sqrt{T_s}} \left(\frac{A_{max}}{A}\right)_1 \quad [9]$$

From continuity considerations, the weight flow to the motor is

$$W_{m1} = W_{s1} - W_{f1} \quad [10]$$

The appropriate substitutions in Equation [10] result in an expression for the volume flow to the motor

$$Q_{m1} = C_1 A_{D0} \frac{R\sqrt{T_s}}{g} \left[\frac{1}{\alpha}\frac{P_s}{P_1}\left(\frac{A_{max}}{A}\right)_1 - (1-\gamma)\right] \quad [11]$$

where

$$\alpha = \frac{A_{D0}}{A_u} \text{ and } \gamma = \frac{x}{x_0} \quad [12]$$

A nondimensionalizing volume flow is defined as

$$Q_{m0} = C_1 \frac{R\sqrt{T_s}}{g} A_{D0} \quad [13]$$

Division of Equation [11] by Equation [13] results in the dimensionless volume flow to the motor

$$q_{m1} = \left(\frac{A_{max}}{A}\right)_1 \left(\frac{1}{\alpha}\right)\frac{P_s}{P_1} - (1-\gamma) \quad [14]$$

Similarly, the expression for the volume flow to the motor for the right-hand three-way valve is

$$q_{m2} = \left(\frac{A_{max}}{A}\right)_2 \left(\frac{P_s}{P_2}\right)\frac{1}{\alpha} - (1+\gamma) \quad [15]$$

Owing to the complexity of the analytical relationships, Equations [14] and [15] were solved by graphical methods with data from reference (2). The assumptions made in the derivation of Equation [14] were tested by an experimental investigation of the weight-flow versus pressure relationships of this type of three-way valve. The results of these tests agree very well with the predicted valve characteristics. Similar agreement for a closed-center three-way valve is reported by Shearer (3).

Four-Way Valve. The steady-state load-flow versus load-pressure characteristics of the four-way valve are obtained by graphically combining Equations [14] and [15]. This operation is accomplished by noting that during steady state

$$q_{m1} = -q_{m2} = q_m \quad [16]$$

The configuration of the four-way valve requires equal values of the variable γ for both three-way valves. The load pressure P_L is

$$P_L = P_1 - P_2 \quad [17]$$

The plot of load flow versus load pressure for a four-way valve with an area ratio $\alpha = 2.0$ is given in Fig. 2.

The pressure sensitivity of a flow valve, defined by

$$\left.\frac{\partial P_L}{\partial \gamma}\right|_{q_m = 0}$$

expresses the ability of the flow valve to furnish a certain load pressure per unit valve displacement with no flow to the motor. The optimum pressure sensitivity of a flow valve is obtained if a supply pressure is applied across the load for an infinitesimal valve opening. The fixed upstream-orifice construction results in a much lower pressure sensitivity than this optimum since one side of the fluid motor retains an appreciable pressure when the other side reaches supply pressure by the closing of its downstream orifice. The pressure sensitivity for valves of different ratios α is given by the inverse slope of the plots in Fig. 3.

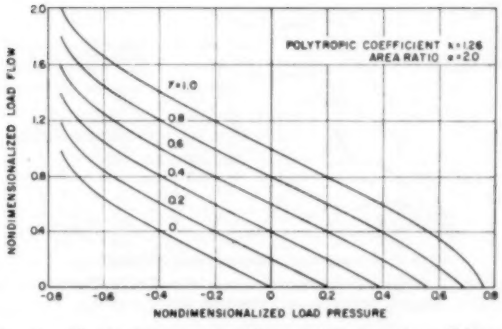


FIG. 2 THEORETICAL LOAD-FLOW VERSUS LOAD-PRESSURE CHARACTERISTICS OF A FOUR-WAY VALVE

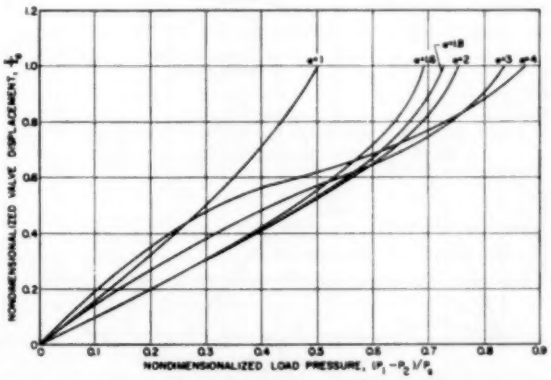


FIG. 3 PLOTS OF LOAD PRESSURE VERSUS VALVE DISPLACEMENT FOR SEVERAL VALUES OF α WITH MOTOR STATIONARY

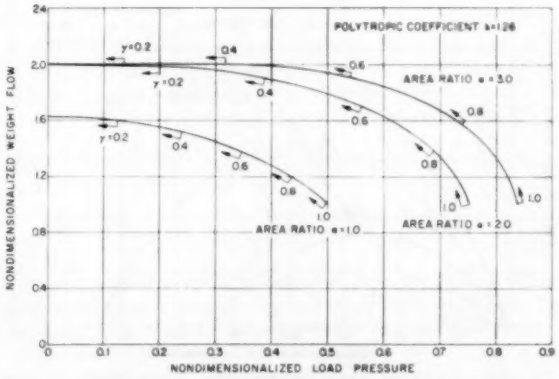


FIG. 4 PLOTS OF WEIGHT FLOW VERSUS LOAD PRESSURE FOR SEVERAL VALUES OF α

The total weight flow to the valve is given by the sum of the weight flows through the upstream orifices

$$W_1 + W_2 = C_1 \frac{P_s}{\sqrt{T_s}} A_u \left[\left(\frac{A_{max}}{A}\right)_1 + \left(\frac{A_{max}}{A}\right)_2\right] \quad [18]$$

Plots of the dimensionless weight flow to the valve

$$(W_1 + W_2)/C_1(P_s/\sqrt{T_s})A_u$$

versus load pressure for several values of α are given in Fig. 4.

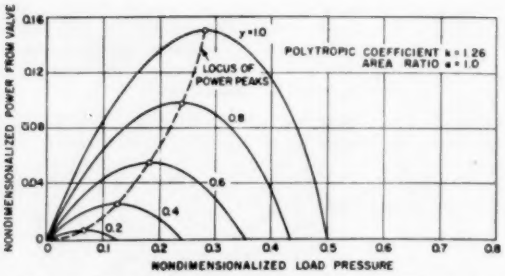


FIG. 5(a) PLOTS OF NONDIMENSIONALIZED FLUID POWER TO THE LOAD VERSUS LOAD PRESSURE FOR $\alpha = 1.0$

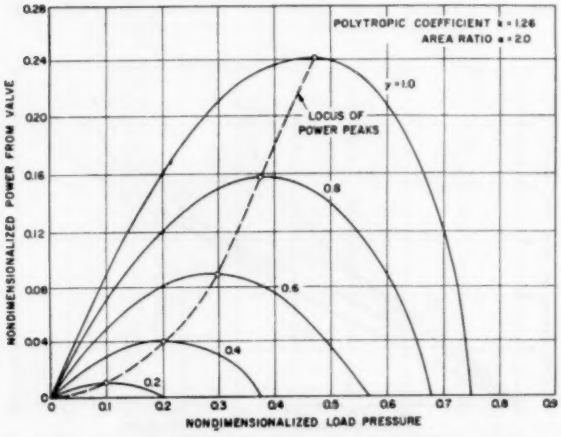


FIG. 5(b) PLOTS OF NONDIMENSIONALIZED FLUID POWER TO THE LOAD VERSUS LOAD PRESSURE FOR $\alpha = 2.0$

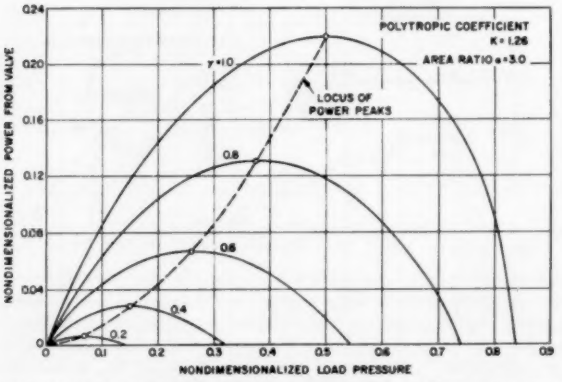


FIG. 5(c) PLOTS OF NONDIMENSIONALIZED FLUID POWER TO THE LOAD VERSUS LOAD PRESSURE FOR $\alpha = 3.0$

Of interest are the facts that, for a fixed value of α, one curve describes the weight flow for all conditions and the weight-flow demand remains remarkably constant over a wide range of operating conditions.

The steady-state output power of the flow valve is given by the product of load pressure and load-volume flow

$$\mathcal{P} = (P_1 - P_2)Q_m \quad [19]$$

Division of Equation [19] by the nondimensionalizing power

$$\mathcal{P}_0 = P_s Q_{m0} \quad [20]$$

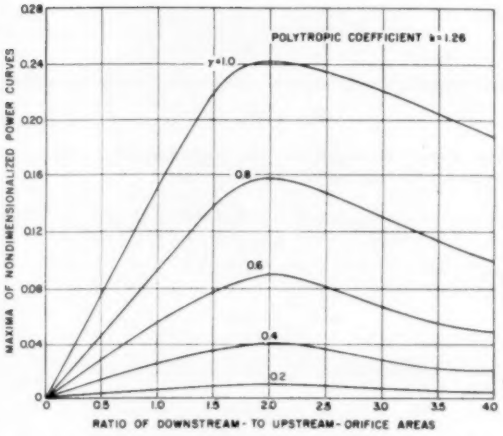


FIG. 6 PLOT OF MAXIMUM POWER VERSUS α FOR SEVERAL VALUES OF γ

results in the nondimensional expression for the power

$$\frac{\mathcal{P}}{\mathcal{P}_0} = \frac{P_1 - P_2}{P_s} q_m \quad [21]$$

Figs. 5(a, b, c) are plots of nondimensional power as expressed in Equation [21] for various values of α. A plot of the peaks of the power curves as a function of α for several values of γ is given in Fig. 6.

A value of $\alpha = 2$ was chosen for the valve on the basis of the following:

1 Linearity of load-flow versus load-pressure characteristics.
2 Optimum pressure sensitivity.
3 Peak power ratio.
4 Compressibility considerations.

The weight-flow curves, Fig. 4, show that the higher values of α result in less variation in weight flow. However, a large value of α will be shown to result in a large compressibility coefficient for the motor.

The choice of α is, therefore, a compromise between compressibility considerations, the nature of the gas generator, and the valve characteristics.

Gas-Motor Analysis. The analysis of motor performance is based on the following assumptions:

1 Perfect gas.
2 No resistance to the flow into the chambers.
3 Adiabatic conditions throughout the charging and discharging processes.
4 Temperature of gas entering the chamber equal to temperature of gases in the chamber.
5 Uniform chamber pressure at all times.

In Fig. 7, the control volume represents an artificial boundary around the chosen volume of gas, across which energy and mass flow. The first law of thermodynamics, when applied to the control volume of Fig. 7, is

$$(u_1 + P_1 v_1)dm - P_1 dV_1 = d(\rho V_1 u_1) \quad [22]$$

The first term represents the energy added to and the work done on the system whenever an incremental mass of gas dm crosses the system boundaries. The second term is the work being done by the system on the piston. The third term is the change in internal energy of the gas within the control volume. Equation

[22] can be modified for use in the gas-motor analysis by substituting the perfect-gas relationships and noting that $dV/dt = \dot{z}A_R$. Equation [22] then becomes

$$\frac{dP_1}{dt} + \frac{kP_1}{V_1}(A_R\dot{z} - Q_{m1}) = 0 \quad \ldots\ldots [23]$$

Equation [23] expresses the action of one side of the fluid motor.

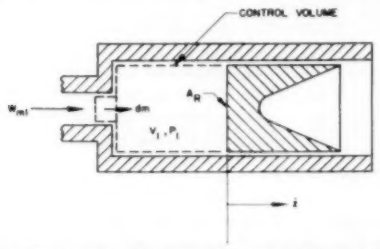


FIG. 7 DIAGRAM OF CONTROL VOLUME IN FLUID MOTOR

Since a positive load flow on one side results in a negative load flow for the other side, the action of the second side of the fluid motor can be expressed by

$$\frac{dP_2}{dt} + \frac{kP_2}{V_2}(-A_R\dot{z} + Q_{m2}) = 0 \quad \ldots\ldots [24]$$

For small load-pressure changes near a no-load condition, the pressures may be expressed as

$$P_1 = P_0 + \Delta P \quad \ldots\ldots [25]$$

$$P_2 = P_0 - \Delta P \quad \ldots\ldots [26]$$

with the result that

$$P_1 + P_2 = 2P_0 \quad \ldots\ldots [27]$$

where the constant no-load cylinder pressure with the valve in mid-position P_0 is given by

$$P_0 = \frac{P_s}{\alpha}\left(\frac{A_{max}}{A}\right)_{\gamma=0} \quad \ldots\ldots [28]$$

Furthermore, operation about the mid-point of the motor travel is assumed and, therefore, volume V_1 remains almost equal to volume V_2

$$V_1 \approx V_2 = V_0 \quad \ldots\ldots [29]$$

Motor action about the mid-position has been shown by an extensive nonlinear analog-computer study (4) to result in the least stable operation.

Combination of Valve and Motor Equations. Substitution of Equations [14], [15], [27], [28], and [29] into the difference of Equations [23] and [24] results in

$$\left(\frac{V_0}{kP_s}\right)\left[\frac{\alpha}{2\left(\frac{A_{max}}{A}\right)_{\gamma=0}}\right]\frac{dP_L}{dt} + A_R\dot{z} = Q_{m0}\left[\gamma - \frac{\alpha}{2\left(\frac{A_{max}}{A}\right)_{\gamma=0}}\left\{\frac{P_L}{P_s} - \frac{2}{\alpha}\left[\left(\frac{A_{max}}{A}\right)_1 - \left(\frac{A_{max}}{A}\right)_2\right]\right\}\right] \quad \ldots\ldots [30]$$

Equation [30] represents the operation of the flow-valve and motor system. The first term represents the compressibility flow necessary for a rate of change in load pressure. The second term is the flow necessary to maintain a steady ram velocity. The right side of the equation represents the flow from the four-way valve and, therefore, is an analytical expression of the valve characteristics. The first term in the outer brackets γ is the spacing, and the second term is the slope of the curves in Fig. 2.

For the case of $\alpha > 1.8$, $(A_{max}/A)_{\gamma=0} = 1$. Furthermore, for the assumed small changes in P_1 and P_2, $(A_{max}/A)_1 = (A_{max}/A)_2 = 1$. With these assumptions, Equation [30] becomes

$$\left(\frac{\alpha}{2}\frac{V_0}{kP_s}\right)\frac{dP_L}{dt} + A_R\dot{z} = Q_{m0}\left(\gamma - \frac{\alpha}{2}\frac{P_L}{P_s}\right) \quad \ldots\ldots [31]$$

The right-hand side of Equation [31] is a linearized expression for the four-way valve steady-state characteristics. The validity of this expression is substantiated by the excellent linearity that exists within the expected operating region of the valve characteristics.

Electromechanical Valve Actuator. The electromechanical transducer was treated as a position and force source having negligible dynamic lag. The force equation of the simplified actuator and its electronic amplifier is given by

$$F_m = K_1K_a e - K_b x \quad \ldots\ldots [32]$$

Since all dynamic effects are neglected, the only force acting on the rotor of the actuator results from the load pressure

$$F_m = P_L A_m \quad \ldots\ldots [33]$$

Gas-Motor Load System. The fluid motor applies a force to the mass m_L in the presence of a load disturbance F_L, damping forces, and a load spring. All damping effects are assumed viscous. The force equation for the load is

$$m_L\ddot{z} + b_L\dot{z} + Kz = P_L A_R - F_L \quad \ldots\ldots [34]$$

Gas Generator. The gas generator is considered a constant-pressure gas source which is unaffected by any system variables.

SYSTEMS ANALYSIS

Open-Loop Characteristics. The open-loop characteristics of the simple system are obtained by combining Equations [31], [32], [33], and [34] with the load-spring stiffness $K = 0$. The resulting differential equation expressing the open-loop system response to an input $e(t)$ in the absence of a load disturbance is given by

$$\left\{(m_L D + b_L)\left[\frac{\alpha V_0}{2kP_s}D + \left(\frac{Q_{m0}}{P_s} + \frac{Q_{m0}A_m}{x_0K_3}\right)\right] + A_R^2\right\}Dz = \left(\frac{Q_{m0}K_1A_RK_4}{x_0K_3}\right)e \quad \ldots\ldots [35]$$

The characteristics of the second-order term in the form of the undamped natural frequency ω_n and the damping ratio ζ are instructive in predicting closed-loop behavior with position feedback. The undamped natural frequency is given by

$$\omega_n^2 = \frac{\frac{2A_R^2kP_s}{\alpha} + kb_LQ_{m0}\left(1 + \frac{A_mP_s}{x_0K_3}\right)}{m_LV_0} \quad \ldots\ldots [36]$$

The damping ratio is given by

$$\zeta = \frac{\frac{b_LV_0\alpha}{2kP_s} + m_L\left(\frac{Q_{m0}}{P_s} + \frac{Q_{m0}A_m}{x_0K_3}\right)}{2\sqrt{\left(A_R^2 + \frac{b_LQ_{m0}}{P_s} + \frac{b_LQ_{m0}A_m}{x_0K_3}\right)\left(\frac{\alpha m_LV_0}{2kP_s}\right)}} \quad \ldots\ldots [37]$$

Therefore, the system damping consists of three independent phenomena:

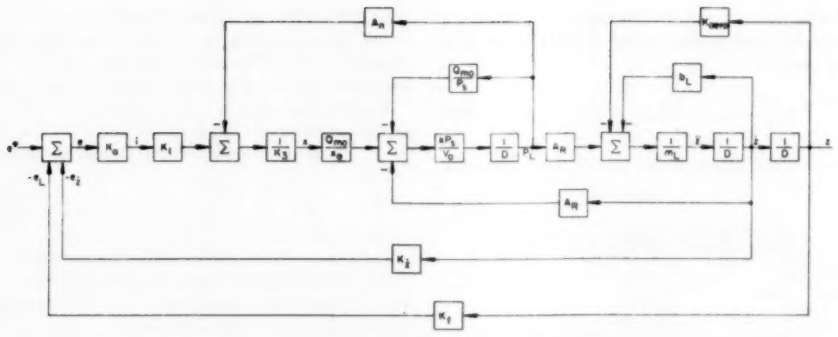


FIG. 8 BLOCK DIAGRAM OF THE CLOSED-LOOP SYSTEM

1 Viscous damping acting on the load.
2 Load-pressure sensitivity of the flow valve.
3 Load-pressure feedback through the flapper valve.

Steady-State Closed-Loop Characteristics. In order for the servomechanism to operate as a position control system, the position of the output member must be compared with the desired position which is in the form of a voltage signal e^*. The load position z, therefore, must be transduced into a voltage e_L. The error voltage is given by

$$e = e^* - e_L \quad [38]$$

The position-transducing constant K_f, which is the position-feedback gain, is defined by

$$e_L = K_f z \quad [39]$$

The steady-state gain of the closed-loop system in the presence of a load spring is given by

$$\frac{\Delta z}{\Delta e^*} = \frac{K_a K_1 P_s A_R}{K(A_a P_s + K_3 x_0) + K_f K_a K_1 P_s A_R} \quad [40]$$

In the absence of the load spring, Equation [40] simplifies to

$$\frac{\Delta z}{\Delta e^*} = \frac{1}{K_f} \quad [41]$$

An important steady-state characteristic of the system is the stiffness of the output member to externally applied forces F_L. The load stiffness is given by

$$\frac{\Delta F_L}{\Delta z} = K + \frac{K_1 K_f A_R P_s}{K_3 x_0 + A_a P_s} \quad [42]$$

The second term represents the internal stiffness of the control system.

Dynamic Closed-Loop Study. The system equations, Equations [31], [32], [33], [34], [38], and [39], can be presented in block-diagram form. A diagram of the third-order closed-loop system is shown in Fig. 8; position feedback and velocity feedback are indicated. The aim of this study is to select values of the system parameters which will result in a satisfactory dynamic response of the system. The well-known graphical techniques of control-system synthesis, when applied to higher than second-order systems, lend themselves best to the study and modification of a single-system variable such as position-feedback gain. Thus every change in any other system constant requires the redrawing of either the root locus or the normalized gain-phase diagrams. The correct feedback gains for a specific response requirement can be determined for each case. The large amount of labor that the graphical techniques entail discourages graphical and analytical experimentation. A further limitation of these methods is the complexity that results if nonlinearities are included. In this study, none of the system constants is fixed. The recourse to machine computing techniques to aid in the final selection of the system constants was, therefore, a natural consequence.

The following example illustrates the design process.

DESIGN EXAMPLE

The performance specifications for a small pneumatic control system are tabulated in Table 1.

TABLE 1 PERFORMANCE SPECIFICATIONS

Steady-state response	
Motor torque at maximum angular velocity, ft-lb	33
Maximum angular velocity, rad/sec	10
Inertia of load, in-lb-sec^2	0.008
External load spring, ft-lb/rad	112
Transient response	
Rise time to a step input (time final amplitude first is reached), millisec	8
Maximum overshoot, per cent	50

Steady-State Performance Calculations. The maximum power output of the system from the data of Table 1 is $\mathcal{P}_{max} = 330$ ft-lb/sec. The nondimensionalized power curves, Fig. 5, for $\alpha = 2$ give a maximum power ratio of $(\mathcal{P}/\mathcal{P}_0)_{max} = 0.24$. The maximum power ratio occurs at a load-pressure ratio of $P_L/P_s = 0.47$. The downstream-orifice area with the valve centered is calculated from the expressions for the nondimensional power, Equations [13], [20], and [21]

$$A_{D0} = \frac{\mathcal{P}_{max} g}{P_s C_1 R \sqrt{T_s} \left(\frac{\mathcal{P}}{\mathcal{P}_0}\right)_{max}} \quad [43]$$

For air at a temperature of 520 F abs and a pressure of 2000 psi, the mid-position downstream-orifice area is $A_{D0} = 0.00107$ sq in. In order to prevent a loss of control at the downstream orifice, the maximum area of the downstream orifice should be no greater than approximately one third of the approach area; therefore the mid-position flow area is one sixth of the approach area, with the result that $D_f \geq 24 x_0$. An assumed orifice opening of 0.003 in. requires a throat diameter of $D_f \geq 0.072$ in. An area ratio $\alpha = 2$ results in the following valve design

$$D_f = 0.113 \text{ in.}$$
$$x_0 = 0.003 \text{ in.}$$
$$A_f = 0.010 \text{ sq in.}$$
$$D_a = 0.026 \text{ in.}$$

For a load-pressure ratio of 0.47, the nondimensionalized motor flow $q_m = 0.52$, Fig. 2. With appropriate substitutions in Equations [13] and [14], the ram flow at maximum motor power is given by

$$Q_m = q_m C_1 \frac{R}{g} \sqrt{T_s} A_{D0} \quad \ldots\ldots\ldots\ldots [44]$$

with the result that $Q_m = 4.30$ cu in/sec and $Q_{m0} = 8.27$ cu in/sec. The load pressure for the maximum power condition is 940 psi. If the radius of the rocker arm is assumed to be 1.20 in., the required ram area is $A_R = 0.421$ sq in. The volume under compression for one side of the fluid motor including all passages is assumed to be 0.30 cu in., which permits an angular deflection of ± 25 deg.

Open-Loop Performance Calculations. The validity of the choice of the system parameters based on the static requirements is examined with the characteristics of the open-loop system. The open-loop undamped natural frequency from Equation [36] is

$$\omega_n\Big|_{b_L=0} = 86.5 \text{ cps}$$

The open-loop damping ratio with no load damping and an estimated stiffness of $K_3 = 4000$ lb/in. for the electromechanical actuator is

$$\zeta\Big|_{b_L=0} = 0.10$$

The undamped natural frequency has a sufficiently high value, but the open-loop damping ratio is too low (5). An appreciable amount of open-loop damping is desirable in order to reduce the effect of variations in the load damping on the stability of the closed-loop system. The system can be stabilized effectively with open-loop damping ratios as low as 0.10, but since only a low value of position-feedback gain can be used in this case, the system becomes very load-sensitive.

A study of Equation [37] reveals that several courses of action may be taken to increase the open-loop damping:

1 The pressure-sensitive area A_a can be increased. If this increase is made and no other system constants are changed, a higher force loading of the valve actuator and a small mid-position orifice opening x_0 result. However, any drastic reduction of x_0 below the chosen value of 0.003 in. is undesirable.

2 The quiescent flow Q_{m0} through the valve can be doubled. If x_0 remains constant, a doubling of Q_{m0} would quadruple A_a. In this manner, the damping ratio ζ is increased to 0.55. However, the power drain on the supply source is doubled.

3 The electromagnetic-actuator net spring stiffness K_3 could be reduced by a factor of 2. However, since a very fast actuator response was assumed in the system analysis, this reduction in the speed of response of the actuator no longer may be neglected.

A compromise solution whereby the upstream-orifice diameter is increased to 0.031 in. and the electromagnetic-actuator stiffness is decreased to 2000 lb/in. yields the following results for the same value of x_0

$$\begin{aligned} Q_{m0} &= 12.0 \text{ cu in/sec} \\ D_f &= 0.165 \text{ in.} \\ A_f &= 0.0211 \text{ sq in.} \\ \zeta &= 0.413 \end{aligned}$$

The effect of these changes on the load sensitivity is negligible because a decrease in the actuator spring stiffness offsets an increase in the load-pressure-sensitive area of the flow valve.

The revised design parameters are not final and are subject to changes during the closed-loop dynamic-response study on the analog computer.

Closed-Loop System Design. The block diagram of the closed-loop system with tachometric feedback is shown in Fig. 8. A high-speed electronic analog computer made by George A. Philbrick Researches, Inc., of Boston, Mass., was used in this design study. The investigation was limited to a study of the linear system. The following assumptions form the basis for the retention of linearity:

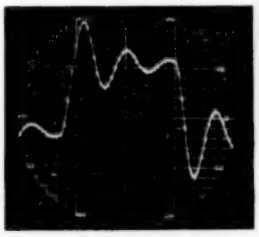
1 Small perturbations are considered in order to simulate linear valve operation.

2 The motor is operated near its mid-position; as a result, the volumes under compression remain essentially equal.

3 The electromagnetic-actuator operation is linear.

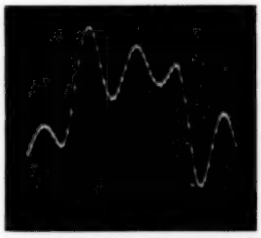
4 All friction phenomena are limited to viscous-type drags.

5 Backlash in the system is negligible.



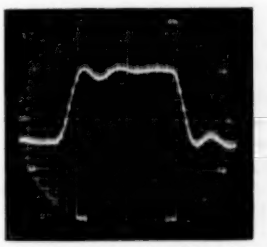
$Q_{mo} = 8.27$ cu in/sec
$D_u = 0.026$ in.
$A_u = 0.00064$ sq in.
$D_f = 0.113$ in.
$A_f = 0.010$ sq in.
$K_3 = 4000$ lb/in.
$A_R = 0.421$ sq in.
$J = 0.008$ lb-sec^2-in.
$V_o = 0.30$ cu in.

FIG. 9 OPEN-LOOP TRANSIENT RESPONSE OF STATICALLY DESIGNED SYSTEM



$K_1K_aK_f = 0.40$
$K_1K_aK_{\dot{z}} = 0$

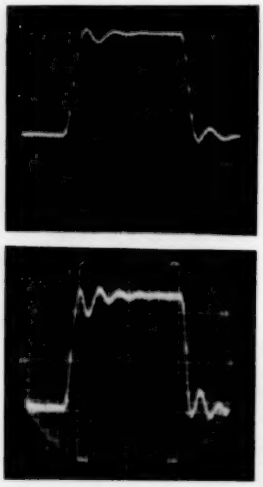
FIG. 10 CLOSED-LOOP TRANSIENT POSITION RESPONSE OF MODIFIED SYSTEM



$K_1K_aK_f = 0.4$
$K_1K_aK_{\dot{z}} = 0.1$
$b_L = 2.0$

FIG. 11 EFFECT OF INTRODUCING VELOCITY FEEDBACK

The accuracy of the analog-computer results was checked by the open-loop velocity response to a step input signal with no load spring. Thus, in Fig. 9, the predicted undamped natural frequency of 86.5 cps and the damping ratio of 0.10 are verified for the unmodified system. An increase in the quiescent flow through the valve and a decrease in the valve-actuator stiffness result in an open-loop velocity response with better damping. The effect of introducing position feedback to the modified system is shown in Fig. 10. The system is oscillatory. The closed-loop position response of the system with a small amount of load damping was effectively stabilized by the introduction of some velocity feedback as shown in Fig. 11. An optimum amount of velocity feed-



a. $K_1K_aK_f = 1.0$
$K_1K_aK_{\dot{z}} = 0.25$
$b_L = 2.0$

b. $K_1K_aK_f = 2.0$
$K_1K_aK_{\dot{z}} = 0.70$
$b_L = 2.0$

FIG. 12 EFFECT OF SIMULTANEOUSLY INCREASING VELOCITY-FEEDBACK AND POSITION-FEEDBACK GAINS

back, based on the least oscillatory system, was found to exist for each value of position-feedback gain and load damping. Furthermore, the adjustment of the velocity-feedback gain for least oscillatory response became increasingly critical with reduced open-loop damping.

The effect of a simultaneous increase in both position-feedback and velocity-feedback gains to attain a faster yet adequately damped system is demonstrated in Figs. 12(*a*, *b*). An attempt was made to increase the feedback gains in such a way that the first peak of the transient response would occur at the final amplitude. The rise time in Fig. 11 is 7 millisec; in Fig. 12(*a*) this has been decreased to 6 millisec, and in Fig. 12(*b*), to 4 millisec.

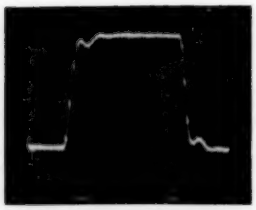
The sensitivity of the transient response of the optimized system to variations in load damping is shown in Figs. 13(*a*, *b*). The effects of doubling and halving the load mass are shown in Figs. 14(*a*, *b*).

The results of the analog-computer study predict that a pneumatic control system with an acceptable response can be designed.

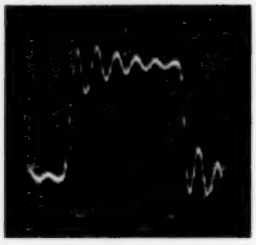
TEST RESULT WITH EXPERIMENTAL MODEL

An experimental model utilizing the dimensions that resulted from the analog-computer study has been built. The design details of the mechanism are illustrated in Fig. 15. The steady-state pressure-sensitivity tests for the flow valve and the load-flow versus load-pressure characteristics showed reasonable agreement with the predicted results. The transient response of the system for small travel of the motor about its mid-position is shown in Fig. 16. The 2000-cps carrier frequency of the position signal indicates that the servomotor has a rise time of about 4 millisec, which is in good agreement with the predicted response. The sinusoidal nature of the overshoot of the position response confirms the assumptions of linear system characteristics, in particular, the viscous load damping resulting from the grease-extruding pistons. Frequency-response tests in the form of Lissajous figures provided further proof that the assumptions of linearity were justified for small motor travel.

Experiments were made with several types of piston seals before the grease-extruding type of piston was devised. Because the analysis and the analog study apply only to viscous friction of the motor, the detrimental effect of stiction could not be predicted. The stabilization of the system with nonviscous friction is far

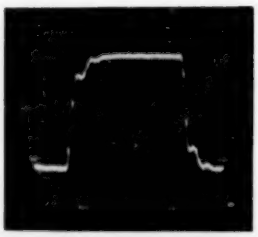


a. $b_L = 4.0$ lb sec/in.

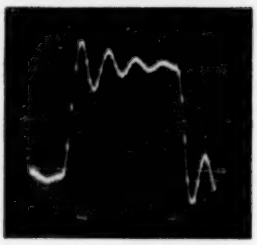


b. $b_L = 1.0$ lb sec/in.

FIG. 13 EFFECT OF VARYING LOAD DAMPING



a. $m_L = 0.0048$ lb-sec^2/in.



b. $m_L = 0.0192$ lb sec^2/in.

FIG. 14 EFFECT OF VARYING LOAD MASS

more difficult, and the resulting motion of the output member occurs in small steps. Owing to the reduced compressibility, the performance of an oil-hydraulic control system was found to be far less sensitive than the equivalent pneumatic system to such nonviscous phenomena. Both rubber and Teflon O-rings, Graphalloy piston rings, and cast-iron piston rings were tested with various boundary lubricants. With the exception of the grease-extruding pistons, semicircular rubber rings placed in undersized grooves with the circular section facing the cylinder wall produced the best results.

CONCLUSIONS

The development of a pneumatic servomotor to meet the exacting requirements of the design example was the result of systematic, thorough preparation. The assumptions used in the linear analysis of the components were checked experimentally. The design of the system was based on the performance predicted by the response studies on the analog computer. Because the effects of variations in system parameters were demonstrated by the analog-computer study, a great deal of time was

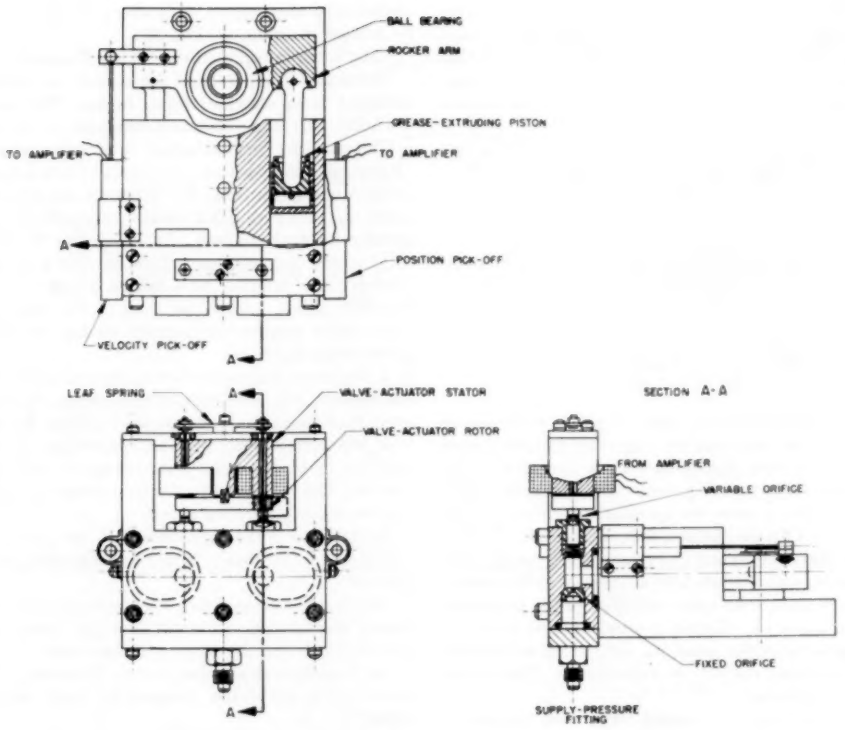


FIG. 15 ASSEMBLY DRAWING OF EXPERIMENTAL MODEL

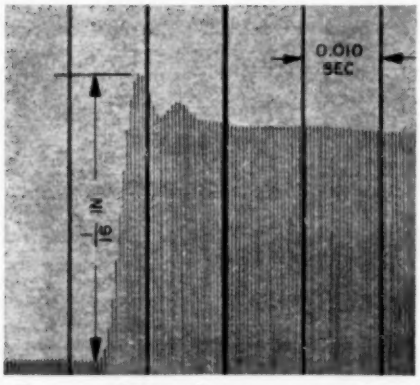


FIG. 16 TRANSIENT POSITION-RESPONSE OF THE EXPERIMENTAL SYSTEM

saved during the process of refining the experimental model. The advisability of designing a pneumatic motor with a minimum of nonviscous friction is an important conclusion that can be drawn from these tests.

To the author, the most gratifying experience of this project was derived from the fact that the system performance could be predicted quantitatively on an analytical basis and with a minimum of cut-and-try fumblings.

ACKNOWLEDGMENT

The author is indebted to Prof. J. A. Hrones and Mr. Emery St. George, Jr., for their guidance in the course of the investigation, to Mr. George Friedensohn for invaluable assistance, and to Mrs. Avis Head for aid in the preparation of this paper.

BIBLIOGRAPHY

1 "The Dynamics and Thermodynamics of Compressible Fluid Flow," by A. H. Shapiro, Ronald Press Co., New York, N. Y., vol. 1, 1953.

2 "Gas Tables," by J. H. Keenan and J. Kaye, John Wiley and Sons, Inc., New York, N. Y., 1948 (revised edition).

3 "Study of Pneumatic Processes in the Continuous Control of Motion With Compressed Air—I," by J. L. Shearer, Trans. ASME, vol. 78, 1956, pp. 233–242.

4 "Continuous Control of Motion With Compressed Air," by J. L. Shearer, M.I.T., Dept. of Mech. Engr., ScD thesis, Cambridge, Mass., 1954.

5 "Dynamic Characteristics of Valve-Controlled Hydraulic Servomotors," by J. L. Shearer, Trans. ASME, vol. 76, 1954, pp. 895–903.

Discussion

C. F. MYER.[4] This paper is an important contribution to the field in that it is the first work to develop and experimentally check system dynamics for a high performance, flapper valve, pneumatic servo.

The use of the flapper-type valve in a high pressure gas system is of particular interest since it requires no lubrication and can handle combustion solids found in propellant powered gas supplies. Although this type of valving has the quiescent power loss associated with open center valving in general, it does, as the author points out, represent a reasonably constant weight flow demand on the power source. This character-

[4] Engineer, Sperry Gyroscope Company, Great Neck, N. Y.

istic, when coupled with a propellant supply that has the characteristic of delivering a constant weight flow, removes the need for a power dissipative pressure regulator. It is my opinion that this paper discloses the existence of the first dependable, high pressure (2000 psi), high performance, pneumatic servomechanism.

In our work, it has been useful to represent the closed-loop, steady-state gain in terms of the internal stiffness and the external spring load. Defining the internal stiffness as in Expression [42]

$$K_s = \frac{K_1 K_a K_f A_R P_s}{K_3 x_0 + A_n P_s}$$

we can write

$$\frac{\Delta Z}{\Delta e^*} = \frac{1}{K_f(1 + K/K_s)}$$

This indicates a general problem in using the gas servo for a positional mechanism where large and varying external spring loads exist; particularly negative spring-like loads (negative hinge moments) that exist when powering some types of missile control surfaces. For example, a servo we investigated yielded a ratio of $K/K_s = \pm 1/2$. This results in a gain change of +100 per cent and —33 per cent for (—) and (+) spring loads, respectively. Since the pneumatic servo will often be an inner-loop of a system with a small gain margin, such gain variations of the pneumatic servo could not be tolerated. Higher internal stiffness can be obtained by using larger amplifier gains, K_a, and/or higher feedback sensitivities, K_f, provided stability is maintained. This is indicated later in this discussion.

The block diagram, Fig. 8, is useful for analog simulation. However, for predictions of response by use of Bode Diagrams and Nichols Charts, Fig. 8 is not too useful. Considering the use of velocity feedback, acceleration feedback which may be of academic use only, and neglecting certain terms which approach magnitudes that negate their effect, we can represent the servo by a block diagram, see Fig. 17.

This, when interpreted on a Bode Diagram, is Fig. 18.

Correction to the straight line curves and phase angles may be obtained from standard servo books. Through the use of a Nichols Chart, the closed loop response can be obtained.

This straight line technique is of value since it permits the analysis of both positive and negative spring-like loads. It also permits one to see that ($K_{\dot{z}}$) depresses the velocity constant and raises ω_n, thus permitting greater values of K_a or K_f resulting in greater internal stiffness. It also shows the reducing effect of ($K_{\dot{z}}$) on the damping ratio of the forward gain transfer function; however, the damping ratio can be adjusted by the acceleration feedback signal. A lead network in the velocity feedback path with corner frequencies properly set can be substituted for the acceleration feedback.

If the servo is designed for a power requirement as outlined by the author, and for a stiffness requirement, then the only parameters that remain to be set are the values of $K_{\dot{z}}$ and $K_{\ddot{z}}$.

It has been our experience in analyzing the linear, sinusoidal problem by both computer studies and by the straight line techniques, that the computer became necessary only for a detailed exploration of a selected design.

In developing the expressions for the power requirements of the servo, I feel that a better understanding is possible if it is pointed out that:

(*a*) Nondimensionalizing volume flow, Expression [13], is the actual steady-state quiescent flow per nozzle measured in the chamber between the orifice and the nozzle.

(*b*) Nondimensionalizing power, Expression [20], is the total quiescent power being dumped by both nozzles for the case where $\alpha = 2$.

A review of Expressions [36], [37], [40], and [42] indicates that for the general case where $\alpha \neq 2$, the term $\alpha/2$ will occur as indicated below. In addition, there appears to be a typographical error in Expression [40] where x_0 was interpreted as K_3

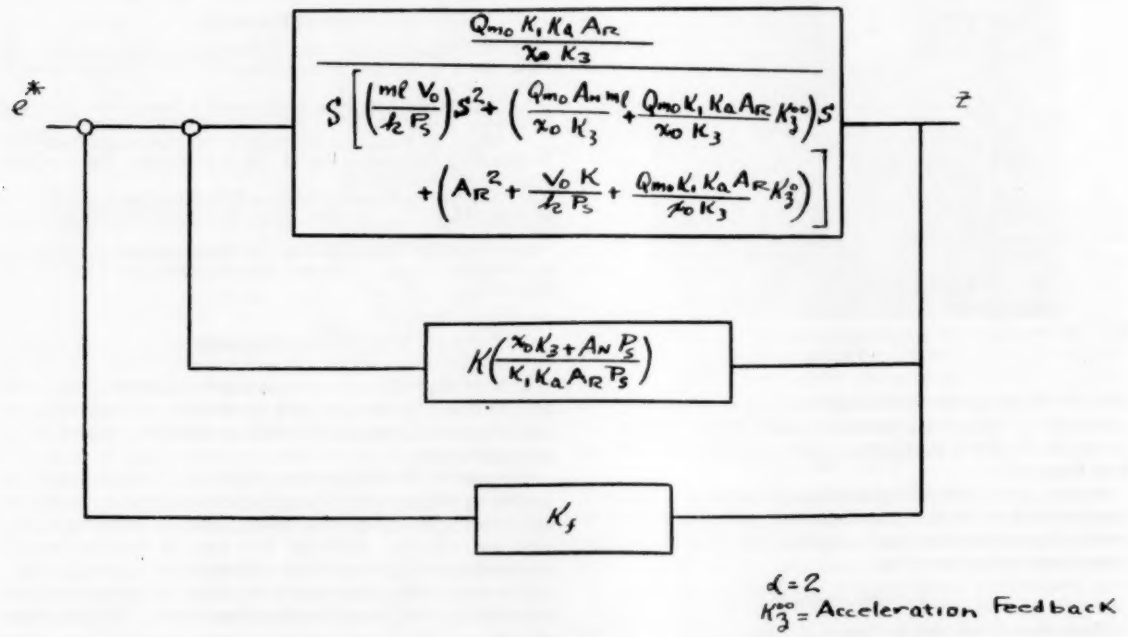


FIG. 17

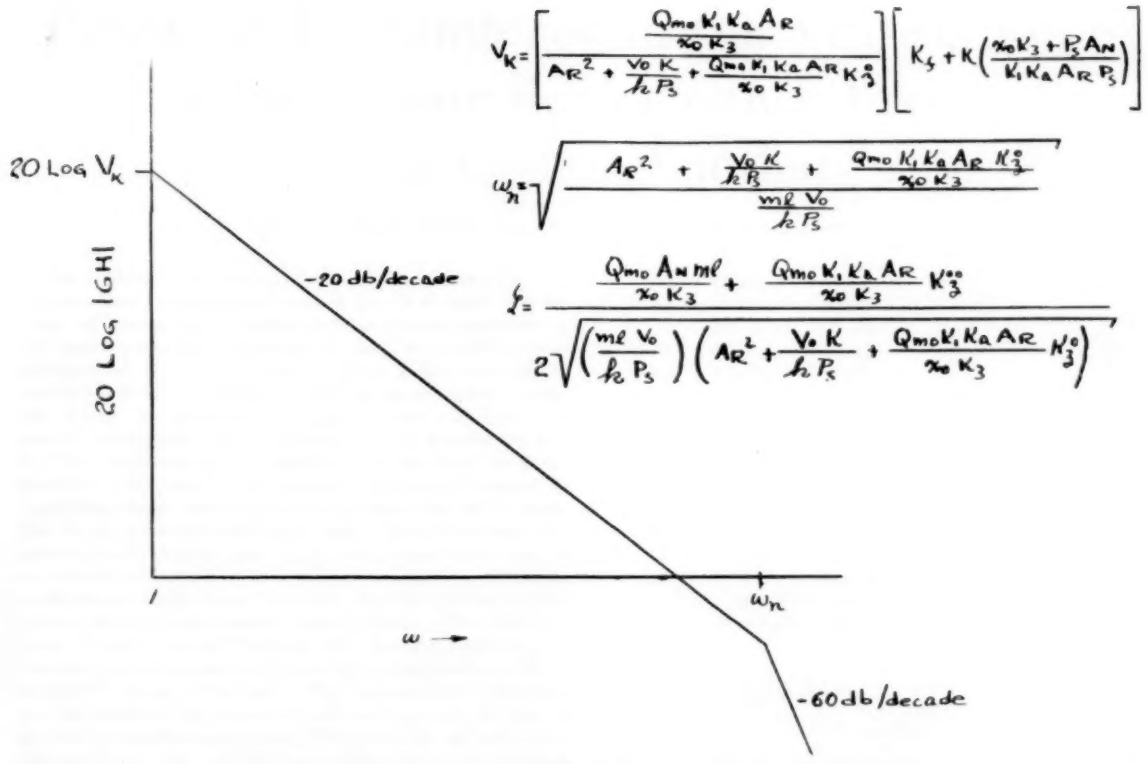


FIG. 18

$$\omega_n^2 = \frac{A_R^2 k P_s \frac{2}{\alpha} + k b_L Q_{m0}\left(1 + \frac{2}{\alpha}\frac{A_n P_s}{x_0 K_3}\right)}{m_L V_0} \quad \ldots [36a]$$

$$\zeta = \frac{\frac{b_L V_0}{k P_s}\frac{\alpha}{2} + m_L\left(\frac{Q_{m0}}{P_s}\frac{\alpha}{2} + \frac{Q_{m0} A_n}{x_0 K_3}\right)}{2\sqrt{\left[\left(\frac{\alpha}{2}\frac{m_L V_0}{k P_s}\right)\left(A_R^2 + \frac{Q_{m0} b_L}{P_s}\frac{\alpha}{2} + \frac{Q_{m0} b_L A_n}{x_0 K_3}\right)\right]}} \quad [37a]$$

$$\frac{\Delta z}{\Delta e^*} = \frac{K_1 K_a A_R P_s}{K(A_n P_s + \frac{\alpha}{2} K_3 x_0) + K_1 K_a K_f A_R P_s} \quad \ldots [40a]$$

$$\frac{\Delta F_L}{\Delta z} = K + \frac{K_1 K_a K_f A_R P_s}{A_n P_s + K_3 x_0 \frac{\alpha}{2}} \quad \ldots [42a]$$

In closing, I should like to emphasize that this paper presents the ground work for design in this field and in this regard we have found the author's work to be of great value in our own work.

AUTHOR'S CLOSURE

The author appreciates Mr. Myer's interesting and constructive discussion of his paper.

The use of such graphical servo synthesis techniques as the Nichols charts, log-modulus diagrams, and the like, is not questioned. However, designing the servo to meet the steady-state power requirements may, as is shown in the paper, result in an underdamped system. In order to obtain an adequate system design, the quiescent power drain may have to be increased appreciably over the statically arrived power level. Thus the power level of the servo valve is in actuality another design variable in addition to the feedback variables K_f, K_i, K_i'', and A_n. It is for this reason that analog computer type of trials are preferred to graphical type of experimentation.

Problems Encountered in the Translation of Compressor Performance From One Gas to Another

By M. J. HARTMANN[1] and W. W. WILCOX,[2] CLEVELAND, OHIO

The problem of translating the performance of a continuous-flow compressor from one gas to another has become of increasing importance from several viewpoints: (*a*) Under certain circumstances it may be possible to test compressors in a special test medium which extends the useful range of test stand to high power or speed ranges. (*b*) It may be desirable to apply a given compressor of known performance to the compression of a different gas. (*c*) The fundametnal information that has been accumulated over the years for air compressors may be useful in improving the performance of compressors for other gases. For single or small numbers of stages the performance for geometrically similar flow at the compressor inlet can be translated rather easily from one gas to another. For large numbers of blade rows, however, the translation of the performance becomes much more difficult. The methods used, however, would indicate the changes necessary to keep the performance level up when a given compressor is applied to some other gas. The experimental methods used in testing compressors in a closed loop using Freon-12 on the test medium are given. The problems of indicating the purity of the test gas and maintaining the desired operating conditions indicate some of the problems compressor test stands using special test medium may encounter. It does appear that special test medium for compressors should be considered as a long-range fundamental study rather than a "quick-fix" extension of existing test facilities.

Nomenclature

The following nomenclature is used in the paper:

A	=	area, sq ft
a	=	velocity of sound, fps
HP	=	horsepower
M	=	Mach number
P	=	total pressure, psf
p	=	stream pressure, psf
R	=	gas constant
ωr	=	element rotational speed, fps
T	=	total temperature, deg R
t	=	stream temperature, deg R
U	=	rotor speed
V	=	stream velocity
W	=	weight flow, lb/sec
$W\sqrt{\theta}/\delta$	=	sea-level equivalent weight flow
β	=	flow angle with axis of rotation
γ	=	ratio of specific heats
η_{ad}	=	adiabatic efficiency
θ	=	ratio of temperature to standard sea-level temperature
ρ	=	density

Subscripts

dg	=	design gas
ideal	=	ideal or no loss condition
tm	=	test medium
0	=	stagnation condition
1	=	rotor-inlet station
2	=	rotor-outlet station

Superscript

$'$	=	relative to rotor

Introduction

In recent years the use of large-capacity continuous-flow compressors has become widespread in the gas turbine-engine field. To obtain accurate aerodynamic performance data over a wide range of flow conditions it is desirable to test the compressor as a separate component rather than in the engine itself. Moreover, in research work it is often necessary to operate at speeds and flow conditions not within existing practice. Requirements for some jet engine-compressor components for example run to 40,000 or 50,000 hp at sea-level inlet conditions.

Numerous suggestions have been made for means of obtaining the experimental performance in gases other than air. Freon-12, for example, has been used successfully in wind-tunnel research and to a limited extent has been used in compressor testing (1).[3] The power requirement in Freon-12 is about 39 per cent of that in air and the rotational speed is about 45 per cent of the air value for equivalent Mach number. Thus the stress level in Freon would be considerably below that for air. Because the ratio of specific heats for Freon-12 is about 1.125, some differences from air must be expected and a means of translating Freon-12 results into air-equivalent performance is necessary. Other combinations of gases have been suggested, as in reference (3), which meet this objection by duplicating the specific-heat ratio of air. Reference (3) indicates that such gas mixtures could reduce power requirements to $^1/_5$ or $^1/_8$ of the air requirement. To the best of the authors' knowledge no test has been run in these "exotic" mixtures. Perhaps the less obvious practical difficulties outweigh the obvious gains.

In recent years a great pool of aerodynamic know-how in the form of empirical and analytical parameters has been built up from compressor tests in air. It would be desirable to apply this air-compressor "know-how" to many industrial problems (for

[1] Head, Section B Compressor Research Branch, Lewis Flight Propulsion Laboratory, National Advisory Committee for Aeronautics.

[2] Project Engineer, Lewis Flight Propulsion Laboratory, National Advisory Committee for Aeronautics.

Contributed by the Gas Turbine Power Division and presented at a joint session of the Gas Turbine Power and Power Divisions at the Semi-Annual Meeting, Cleveland, Ohio, June 17–21, 1956, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, May 17, 1956. Paper No. 56—SA-61.

[3] Numbers in parentheses refer to the Bibliography at the end of the paper.

example, pumping natural gas, cooling gases, and many industrial processes). As these principles are applied to compressors dealing with less familiar gases it should be possible to improve the level of performance of these compressors. Again a means of translating test results from one gas to another is required.

In this paper the methods of translating compressor performance from gas to gas will be discussed. Some examples are given using air as design gas and Freon-12 as the test medium. Limitations of the system are discussed. The experimental methods and the equipment used in a compressor-test rig using Freon-12 as a test medium are given. In addition, the problems encountered are discussed as they may affect the practicablity of using Freon-12 or other gases as a test medium.

Compressor Performance in Gases With Different Ratios of Specific Heat

This section will deal with the problems of converting compressor performance from one gas to another in which the ratio of specific heats is different. The general approach will be to set operational conditions equivalent to obtain geometrically similar inlet flow and then to evaluate the changes in over-all flow conditions and performance. It is assumed that both gases can be described with sufficient accuracy by the perfect gas laws.

Methods of Setting Equivalent Speed and Weight Flow. Continuous-flow compressors such as the axial-flow and the mixed-flow compressor rotors, shown in Fig. 1, will be considered in this paper. It can be presumed that aerodynamic performance will be similar when geometrically similar flow is obtained everywhere around the rotor blades and passages. Geometric similarity is obtained when the flow angles, Mach numbers, and the physical geometry of the compressors are the same. This condition of geometric similarity is probably sufficient since for this type of compressor the sizes and flow velocities are of sufficient magnitude so that the Reynolds number on all bounding surfaces does not change from a turbulent to a laminar boundary condition and the effects of compressiblity are of major importance. A more complete discussion of these conditions of similarity is given in reference (2).

To obtain geometrically similar flow at the compressor-rotor entrance it is necessary that the inlet relative flow angle and Mach number be set equal. Typical entrance Mach-number diagrams are shown in Fig. 2. These will be referred to as equivalent entrance Mach-number diagrams (since the flow angles and Mach numbers are equivalent). Equivalent entrance diagrams will be obtained in different test mediums when the rotational speeds are set by Equation [1]. For identification purposes one gas has been

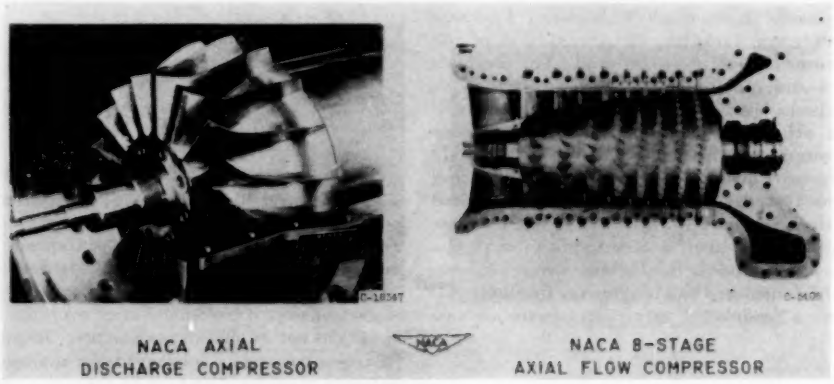


Fig. 1 Continuous-Flow Compressors Investigated by the NACA

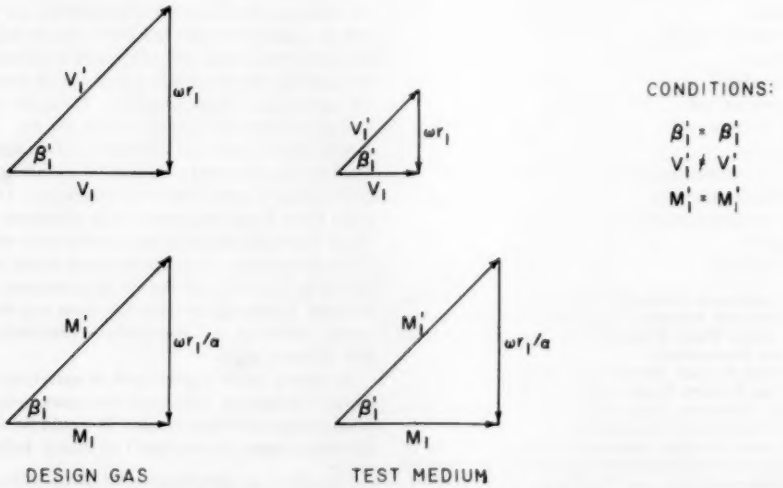


Fig. 2 Inlet Velocity and Mach Diagrams of Compressor Rotor for Design Gas and Test Medium

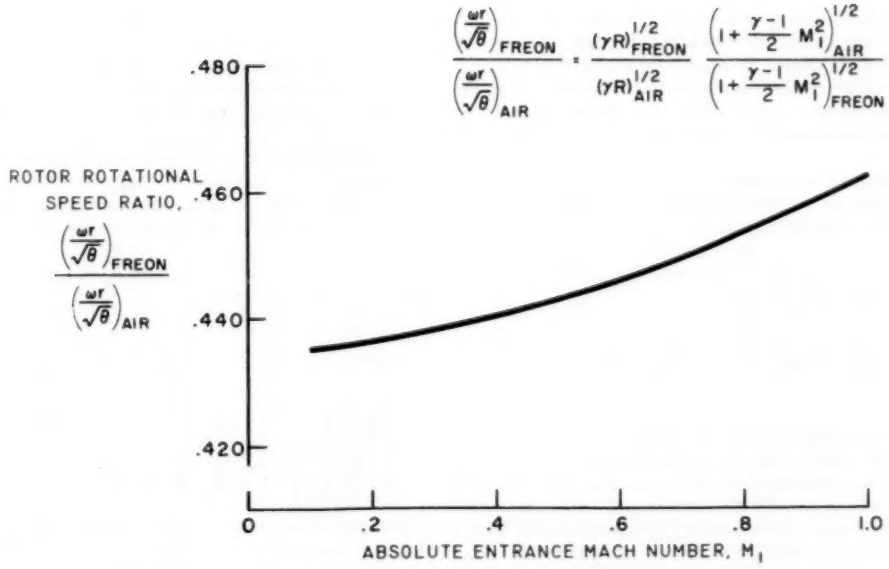


FIG. 3 RELATION OF ROTATIONAL SPEED IN FREON-12 AND AIR

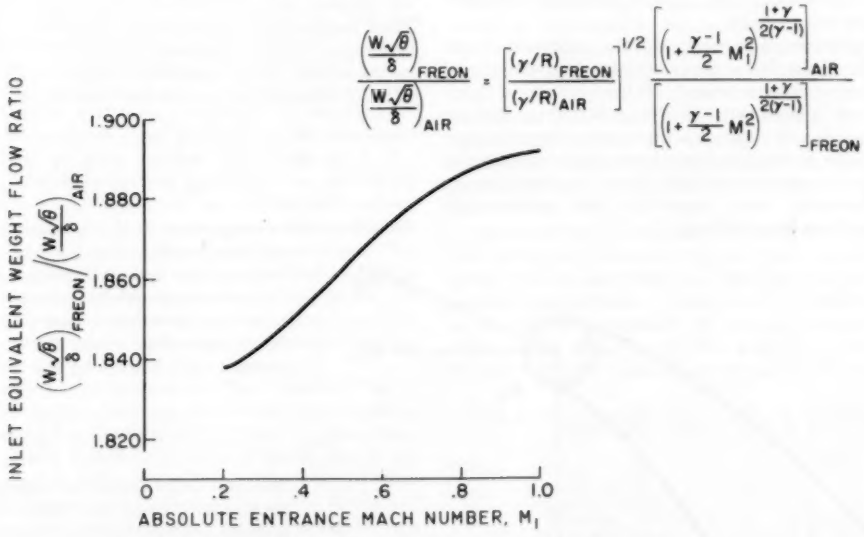


FIG. 4 RELATION OF WEIGHT FLOW OF FREON-12 AND AIR

referred to as the design gas (dg) and the other as the test medium (tm)

$$\frac{\left(\frac{U}{\sqrt{\theta}}\right)_{tm}}{\left(\frac{U}{\sqrt{\theta}}\right)_{dg}} = \left[\frac{(\gamma R)_{tm}}{(\gamma R)_{dg}}\right]^{1/2} \left[\frac{\left(1+\frac{\gamma-1}{2}M_1^2\right)_{dg}}{\left(1+\frac{\gamma-1}{2}M_1^2\right)_{tm}}\right]^{1/2} \quad \ldots [1]$$

(All equations are developed in the Appendix.) Thus, when the absolute inlet Mach number is known and is the same in both gases, the ratio of rotational speeds in the two gases to obtain geometrically similar flow at the compressor inlet can be determined by Equation [1]. It can be seen readily that if a radial variation in inlet absolute Mach number exists over the inlet annulus (when inlet guide vanes are used, varying inlet Mach numbers is a general case), the ratio of rotational speeds for obtaining the matched inlet diagram would be different for various elements along the inlet radius. Conversely, with a fixed speed ratio, and variable M_1, it is impossible to obtain exactly geometrically similar flow over the entire compressor inlet. The error involved may be judged by the variation of the second term on the right-hand side of Equation [1] as the absolute value of the inlet Mach number varies.

In Fig. 3 the ratio of rotational speeds for Freon-12 to air for various absolute inlet Mach numbers is shown. (Freon-12 in this superheated region departs somewhat from a perfect gas but

for this comparison an average $\gamma = 1.125$ was used, reference 3.) The ratio of rotational speeds varies from about 0.435 to about 0.462 over the range given. It can be presumed from this that small variations in inlet absolute Mach number along the radius (usually $M_{1,max} - M_{1,min} < 0.1$) would result in only a small departure from geometrically similar flow along the compressor-inlet radius.

For the foregoing conditions of equivalent entrance Mach-number diagrams, the ratio of equivalent weight flow between the two gases can be given as

$$\frac{\left(\frac{W\sqrt{\theta}}{\delta}\right)_{tm}}{\left(\frac{W\sqrt{\theta}}{\delta}\right)_{dg}} = \left[\frac{\left(\frac{\gamma}{R}\right)_{tm}}{\left(\frac{\gamma}{R}\right)_{dg}}\right]^{1/2} \frac{\left[\left(1 + \frac{\gamma - 1}{2} M_1^2\right)^{\frac{1+\gamma}{2(\gamma-1)}}\right]_{dg}}{\left[\left(1 + \frac{\gamma - 1}{2} M_1^2\right)^{\frac{1+\gamma}{2(\gamma-1)}}\right]_{tm}} \quad \ldots\ldots [2]$$

Similarly to the ratio of rotational speeds, the ratio of equivalent weight flows between the two gases is dependent only on the gas properties and absolute inlet Mach number. The ratio of weight flows is a constant only when M_1 is constant over the entire radius of the inlet annulus. Equation [2] has been solved over a range of inlet absolute Mach numbers for Freon-12 and air and is plotted in Fig. 4. The ratio of weight flows varies from about 1.84 to 1.89 over the range given. Thus, for small gradients of inlet Mach number ($M_{1,max} - M_{1,min}$ small) the weight-flow ratio will be considered as a constant.

The preceding discussion has dealt with the problem of obtaining geometrically similar flow at the compressor-rotor entrance for a change in the gas being compressed. It has been shown that if the absolute Mach number varies over the radius of the inlet annulus it is not possible to obtain exact geometrically similar flow all along the radius of the compressor inlet at a given operating condition. For the gradients of inlet Mach number normally encountered, however, these departures from geometrically similar flow would not be significant.

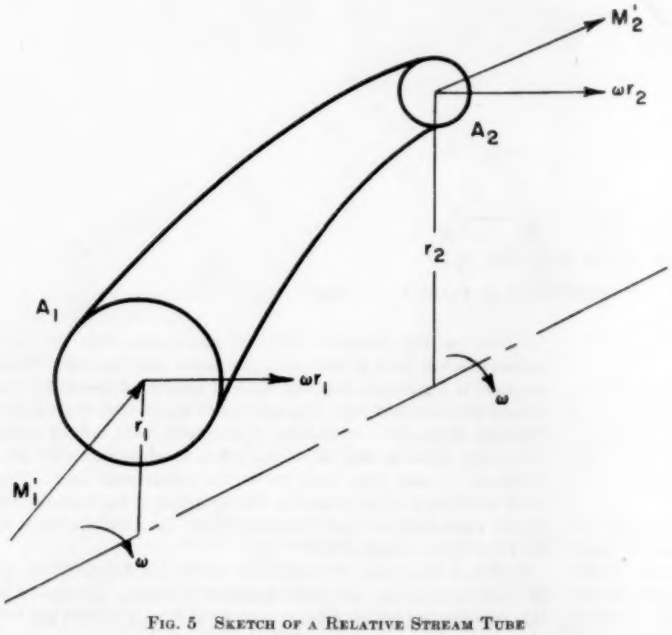


FIG. 5 SKETCH OF A RELATIVE STREAM TUBE

Relative Flow Through the Rotor Passage. The preceding section has indicated a method of obtaining geometrically similar flow over the rotor entrance and the necessary departures from similarity between gases with different ratios of specific heat. We may now examine the flow through the rotating passage. Since geometric areas remain fixed it will be assumed that the effective flow-area ratio of a relative stream tube remains the same for the two gases. Such a relative stream tube is sketched in Fig. 5. The flow Mach number in the two gases at the discharge from the rotor can be related by the continuity equation in the form

$$M_{2,tm}' \left(1 + \frac{\gamma - 1}{2} M_{2}'^2\right)_{tm}^{\left(\frac{\gamma+1}{2(\gamma-1)}\right)_{tm}} = \left[\frac{\left(\frac{T_1'}{T_2'}\right)_{dg}}{\left(\frac{T_1'}{T_2'}\right)_{tm}}\right]^{1/2} \frac{\left(\frac{P_2'}{P_1'}\right)_{dg}}{\left(\frac{P_2'}{P_1'}\right)_{tm}} \frac{\left[\left(\frac{1 + \frac{\gamma - 1}{2} M_2'^2}{1 + \frac{\gamma - 1}{2} M_1'^2}\right)^{-\frac{\gamma+1}{2(\gamma-1)}}\right]_{dg}}{\left[\left(\frac{1}{1 + \frac{\gamma - 1}{2} M_1'^2}\right)^{-\frac{\gamma+1}{2(\gamma-1)}}\right]_{tm}} \quad \ldots\ldots [3]$$

The relative inlet and outlet Mach number to the stream tube are assumed or known for the design gas. The relative inlet Mach number in the test medium is assumed equal to that in the design gas to obtain geometrical similarity. Without heat transfer the relative total temperature change can be considered as an effect of change in radius over the rotor and is given by

$$\frac{T_1'}{T_2'} = \frac{1}{1 + \frac{\omega r_2^2}{a_{0,1}'^2} \frac{\gamma - 1}{2} \left[1 - \left(\frac{r_1}{r_2}\right)^2\right]} \quad \ldots\ldots [4]$$

Thus the relative temperature ratio is a function of the geometry and inlet conditions of gas properties and can be evaluated for both gases being considered.

However, two effects must be considered in determining the relative total pressure ratio along the stream tube. First, the ideal relative total pressure ratio is given by the temperature ratio

$$\left(\frac{P_2'}{P_1'}\right)_{ideal} = \left(\frac{T_2'}{T_1'}\right)^{\frac{\gamma}{\gamma-1}} \quad \ldots\ldots [5]$$

This may be evaluated for both gases. However, for flow with losses, the relative total pressure P_2' is somewhat less than the ideal relative total pressure ratio $P_{2,ideal}'$ and may be expressed as

$$\frac{P_2'}{P_1'} = \left(\frac{P_2'}{P_1'}\right)_{ideal} \left(\frac{P_2'}{P_{2,ideal}'}\right)$$

Usually the loss in total pressure is given as a fractional value of the inlet dynamic head and it is probably a reasonable assumption that this loss is equal in the two gases. This may be expressed as

$$\left(\frac{P_{2,ideal}' - P_2'}{P_1' - p_1}\right)_{tm} = \left(\frac{P_{2,ideal}' - P_2'}{P_1' - p_1}\right)_{dg}$$

It will be presumed that this loss in dynamic head is known for the design conditions; that is, it may have been assumed in the design or may have

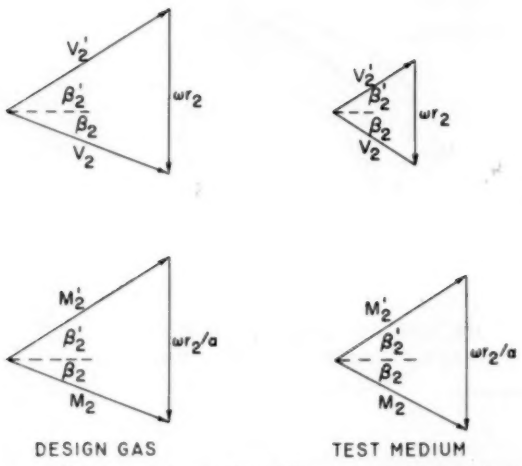


FIG. 6 DISCHARGE FLOW DIAGRAMS IN DESIGN GAS AND TEST MEDIUM

been obtained from a design adiabatic efficiency or experimental data. Then the foregoing expression can be used to obtain the actual relative total pressure $(P_2')_{tm}$ test medium, since this is the only unknown. Thus all necessary values can be obtained to evaluate the right-hand side of Equation [3]. The relative discharge Mach number can then be determined. In this case $(\mathbf{M}_2')_{tm}$ was determined by plotting values of the left-hand side of the equation for the desired range of Mach numbers.

By working from absolute to relative inlet conditions and from inlet to outlet relative conditions the relative temperature and pressure and Mach number at the discharge can be determined. The relative discharge Mach number can be converted to a velocity by obtaining the velocity of sound from the relative discharge temperature. Now presuming that the relative discharge flow angle is the same for both gases and is known, the discharge-vector diagram shown in Fig. 6 can be obtained. The rotor-performance values (absolute total pressure ratio and total temperature ratio) and efficiency can be determined in the test medium. Thus a method of translating the compressor performance along a stream tube from one gas to another has been outlined.

Relation of Analytical Performance and Flow Variables. To obtain a feel for the change in performance from one gas to another Table 1 was prepared to illustrate the relative conditions for a series of five arbitrary vector diagrams in Freon-12 and in air. In all cases the absolute inflow direction is axial. The first three columns of this table are for a relative inlet angle of 45 deg and relative discharge flow angle of 35 deg (all angles measured from the axis of rotation) and the inlet and discharge axial velocities are set equal. The rotational speeds are, respectively, 800, 900, and 1000 fps in air. Columns 4 and 5 are for vector diagrams of 60 deg relative inlet flow angle and 0 deg relative outlet flow angle and for rotational speeds of 1100 and 1300 fps. The change in relative stream conditions over the second two rotors would be rather large. The five vector diagrams thus computed can be considered as covering a very wide range of conditions. The following important observations can be made from this table:

1 The ideal total pressure ratio is shown as the ratio of total pressure ratio of Freon-12 over that for air. In the three vector diagrams of case I this ratio is slightly less than I. In case II where the change in relative stream conditions is rather large the effect of changing gases on total pressure ratio is somewhat larger.

2 The second point of interest here is the fact that if an arbitrary but equal loss is applied in Freon and in air the efficiency in Freon is somewhat higher than in air.

3 The ratio of power required in Freon to that of air is about 0.377 to 0.399.

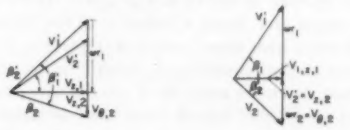
4 If additional blade rows are used such as stators or additional stages it is important to know the change in absolute flow Mach number and absolute flow angle between the two gases. In general the absolute discharge Mach number from these arbitrary vector diagrams is slightly lower in Freon than in air. Again in the latter two cases a sizable difference in Mach number does exist. The discharge absolute flow angle is somewhat larger for Freon than for air. Whereas this amounts to only 1.54 deg at the greatest point it should be pointed out that such a change in flow angle to following blade rows may cause a sizable change in flow conditions, especially if it is compounded through several stages.

5 The ratio of weight flows and rotational speed are tabulated in this table. The weight-flow ratios range from 1.87 to 1.89 and the rotational-speed ratios range from 0.445 to 0.46.

These five arbitrary vector diagrams serve to illustrate the variations in computed performance between Freon and air obtained when geometrically similar flow at the rotor entrance is set and constant area ratios are assumed over the rotating channel. The methods outlined do indicate that there would be some differences in performance between Freon and air. The following section will consider the problems encountered when these methods are applied to an actual compressor rotor and will compare a few test results obtained in Freon to those obtained in air.

Application of Transformation Methods to a Compressor Rotor. The methods of translating for the single stream tube, the compressor performance outlined in the preceding section can be applied to a rotor passage. Approximate translation could be made of the over-all performance by using a mass-average inlet and outlet vector diagram. In this case the relative stream-tube changes would be taken as an average of the changes over the

TABLE 1 ANALYTICAL RELATION OF COMPRESSOR ROTOR PERFORMANCE IN FREON AND AIR



DESIGNATION	SYMBOLS	CASE I			CASE II	
ASSUMED ROTATIONAL SPEED IN AIR, RPM	$(\omega r/\sqrt{\theta})_{AIR}$	800	900	1000	1100	1300
RATIO OF PRESSURE RATIOS	$\frac{(P_2/P_1)_{FREON}}{P_2'/P_1')_{AIR}}$	.982	.983	.983	.93	.96
RATIO OF EFFICIENCIES	$\frac{\eta_{AD,FREON}}{\eta_{AD,AIR}}$	1.01	1.01	1.06	1.06	1.03
RATIO OF HORSEPOWERS	$\frac{HP_{FREON}}{HP_{AIR}}$	.377	.384	.399	.375	.382
CHANGE IN DISCHARGE MACH NUMBER	$M_{2,FREON} - M_{2,AIR}$	-.001	-.001	-.004	-.041	-.073
CHANGE IN DISCHARGE FLOW ANGLE	$\beta_{2,FREON} - \beta_{2,AIR}$	.19	.39	.49	1.25	1.54
WEIGHT-FLOW RATIO	$\frac{(W\sqrt{\theta}/b)_{FREON}}{(W\sqrt{\theta}/b)_{AIR}}$	1.88	1.88	1.89	1.87	1.88
RATIO OF ROTATIONAL SPEEDS	$\frac{(\omega r\sqrt{\theta})_{FREON}}{(\omega r\sqrt{\theta})_{AIR}}$	.45	.45	.46	.44	.458

(A) TOTAL PASSAGE AREA AS A SINGLE ROTATING STREAM TUBE

(B) SERIES OF STREAM TUBES WHICH TOGETHER FORM THE TOTAL PASSAGE

FIG. 7 GAS CONVERSION

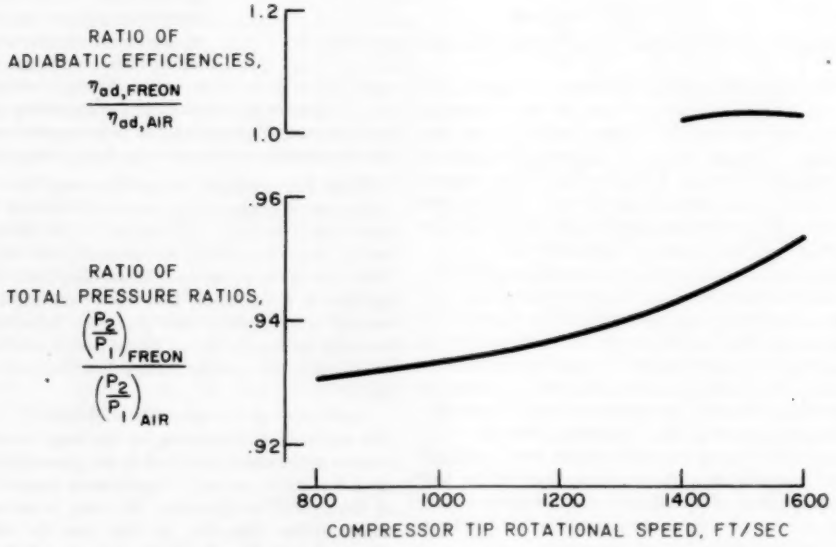


FIG. 8 SUPERSONIC-COMPRESSOR PERFORMANCE IN AIR AND FREON-12 FROM REFERENCES (1, 4)

total rotor passage as shown in Fig. 7(A). If the distribution of discharge angle and Mach number over the discharge annulus must be known with some precision (especially if the rotor-discharge annulus is of considerable height), it may be desirable to make a more extensive analysis of the discharge-flow conditions. It would be necessary under these circumstances to divide the rotor passage into a number of stream tubes and compute the changes along each of these stream tubes as illustrated in Fig. 7(B). Thus the means of translation may vary slightly from one case to another depending on the requirements.

The following discussion will consider some comparisons of rotor performance in Freon and in air for which the inlet conditions are set to give approximate geometric similarity. The data for references (1) and (4) are for two rotors of different sizes but of the same geometric design. (Area ratio and angles were maintained equal.) The rotor of reference (1) was 16 in. diam and run in Freon, whereas reference (4) was 24 in. diam and was run in air. A few comparisons of peak total pressure ratio and adiabatic efficiency are indicated in Fig. 8. Over the range of speed given the total pressure ratio in Freon ranges from about 93 to a little over 95 per cent of that obtained in air, and the efficiency is between 2 and 4 percentage points higher in Freon than in air. These differences in performance between Freon and air are of the order indicated by the arbitrary vector diagrams used in the preceding section.

Another comparison was obtained from unpublished data where a few data points were obtained in Freon and in air. One data point at which matching inlet conditions were nearly established is given in Fig. 9. The relative Mach number in the two gases and the over-all performance are tabulated above the figure. The rotor-inlet relative Mach number in air was 0.02 less than Freon, thus there is a slight departure from matching the inlet Mach-number condition. The ratio of mass-averaged over-all total pressure ratio obtained in Freon was about 0.99 of that obtained in air and the adiabatic efficiencies were equal in Freon and in air. Of special interest in this figure is a variation of pressure ratio, Mach number, absolute discharge, and absolute discharge angle across the discharge, passage. At the hub the total pressure ratio was slightly lower in Freon than in air, whereas at the tip the total pressure ratios are equal. The absolute discharge Mach

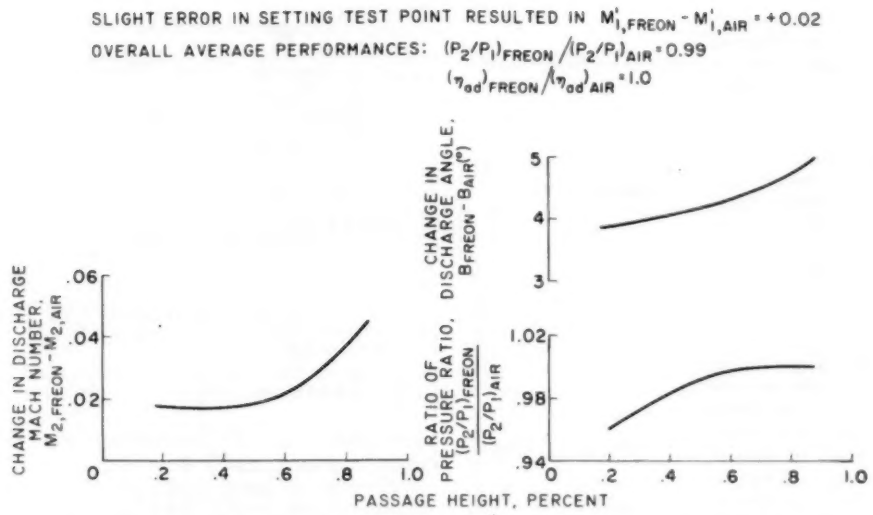


FIG. 9 RELATION OF FLOW CONDITIONS BEHIND A COMPRESSOR ROTOR IN FREON-12 AND AIR

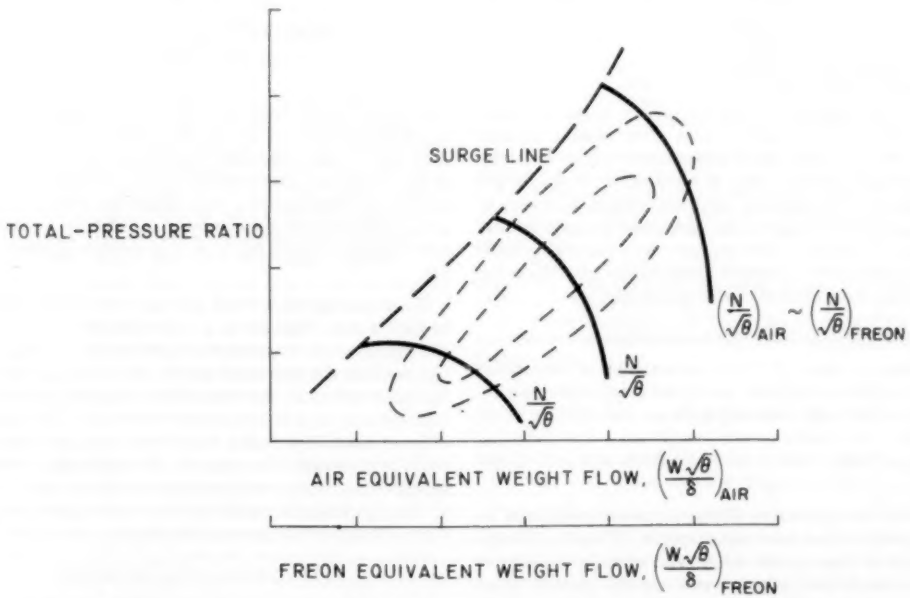


FIG. 10 COMPRESSOR PERFORMANCE MAP SHOWING APPROXIMATE CONVERSION FROM AIR TO FREON-12

number indicated is slightly higher for Freon than in air. This is contrary to the arbitrary vector diagrams shown in the previous section. However, the difference is small and may be in the order of inlet mismatching between the Freon and air data points. The absolute discharge flow angle seems to be about 4 to 5 deg higher in Freon than in air. This is the direction of change indicated by the arbitrary vector diagrams; however, the change is somewhat larger than that which may have been expected. Here again the slight mismatching between Freon and air at these particular points may have caused a portion of the larger difference in discharge flow angle.

It has been indicated by the arbitrary vectors and the few bits of data available that, if geometrically similar flow was set, the pressure ratio and adiabatic efficiency for a single-blade row are nearly the same for Freon and air. Thus, if the performance map of a compressor is available as shown in Fig. 9 for one gas, the performance may be estimated in another gas if the equivalent weight flow and rotational speed are changed to the required values. To make these changes it is only necessary to know the gas properties and the approximate absolute inlet Mach number. The performance map may then be considered as an approximate map in the other gas since the total pressure ratio and adiabatic efficiency would be essentially unchanged. This could be verified by a few conversions at specific points as outlined in the preceding

sections. It should be pointed out that the arbitrary vector diagrams used and the experimental performances given were all for a condition of little or no radius change over the rotor ($r_1/r_2 = 1.0$). In the case of large changes in radius such a simple translation of compressor performance probably would have to consider this effect on temperature and pressure rise.

An entirely different condition exists, however, in the case of a multistage compressor. Geometrically similar flow could be set up in the first stage. However, the succeeding stages would then be operating away from their design point; that is, the Mach number and flow angle would be changed from their design. Thus the change of gas for this type of compressor becomes very complicated and requires additional information. It would be necessary to have available data or the necessary assumptions for the slightly different flow angles and Mach numbers. This deviation in flow conditions may not become serious for several stages and thus the methods used may be useful. Also after passing through several blade rows to obtain good performance it may be necessary to set the blades at new angles to account for these changes in flow conditions. The method used here would be useful to determine the desirable changes. In such cases it would be necessary to consider the detail flow changes over stationary blade rows as well as the rotors. These changes in flow conditions can be handled by the same method used in this paper.

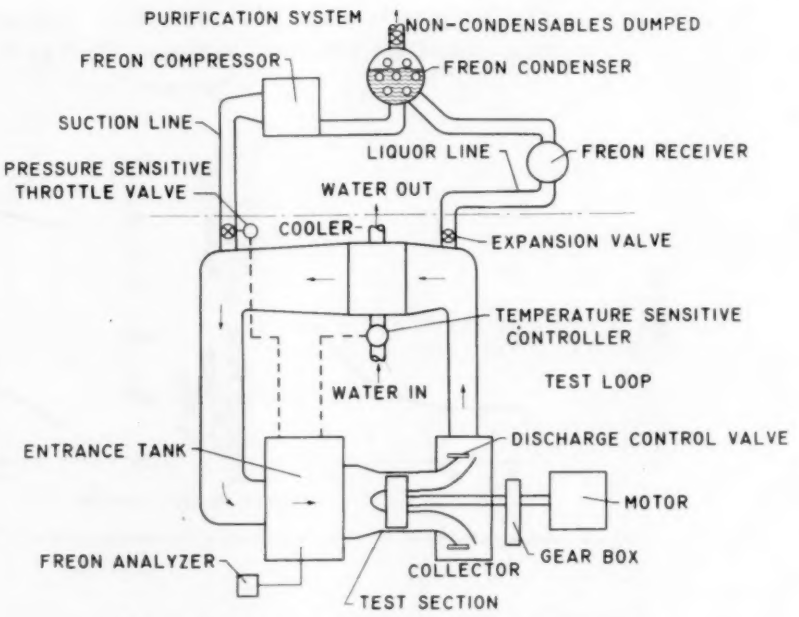


FIG. 11 FLOW SHEET OF FREON TESTING OF A COMPRESSOR ROTOR

THE USE OF SPECIAL TEST MEDIUM FOR COMPRESSORS

The approximate nature of the transformation of compressor performance for gases of different specific-heat ratios has been indicated. This would imply that such a gas as a test medium would find application only under certain conditions. The following are examples in which it may be desirable to use such a test medium:

1 Extremely high equivalent effective rotative speeds may be obtained at relatively low mechanical speeds. Thus the aerodynamic problems at these speeds can be studied independently of the mechanical problems and thus evaluate the possible gains available.

2 If model manufacturing cost or time is greatly reduced because of the reduced stress level at the lower rotative speeds, the research program may be facilitated in this manner.

3 Testing in such a medium should probably be limited to a small number of stages since in general the problem of matching flow angles and velocities into additional blade rows will become serious as the number of stages is increased.

4 It would seem that testing in a medium with a change in ratio of specific heats should be limited to investigating operational principles since the approximate nature of the results would not lend itself to obtaining design control data. As design methods are improved and reduced to basic flow parameters this limitation may be partially removed.

5 A long-range research program in which the models are designed for the test medium could obtain considerable design data at relatively small investment in power and test equipment. The data must then be reduced to basic flow parameters to be useful in the design of compressors for the required gas. This would be a possible advantage in a very long-range program. Such a program would require development and the construction of special test models. This is the type of program suggested in reference (4).

Using gas mixtures which give a specific-heat ratio equivalent to that of the design gas as a test medium would avoid some of the aerodynamic translation of performance values. In such a test medium the rotational speeds are related by the functions of the gas constant R; however, there is a great reduction in power requirements as indicated in reference (3). The proper mixture of gases must be supplied to a closed test loop and closely controlled to maintain this mixture. It is thought that some of the mixture and control problems may be indicated by a discussion of the experimental problems that were encountered in testing Freon-12 and will be given in the following section.

EXPERIMENTAL PROBLEMS

A schematic diagram of the test rig using Freon-12 as a test medium for a compressor investigation is shown in Fig. 11. The test loop is closed. The test medium flows from the entrance tank through the compressor rotor and test section, the collector, the coolers, and back to the entrance tank. One throttle valve was used at the discharge of the test section. The back pressure on the compressor rotor is set by this throttle valve. The design is such that a minimum volume of high-pressure test medium is maintained between the compressor rotor and the throttle valve (thus the quantity of Freon in the test loop is kept small). The remainder of the test loop was maintained near atmospheric pressure. The pressure level in the entrance tank was maintained near atmospheric pressure by the quantity of gas in the test loop. In spite of the fact that the seals around the shaft were reasonably good and the usual pipe flanges were used with gaskets and ex-

tensive efforts were made to seal all joints, considerable leakage both in and out of the test loop always occurred. Since there is a vast number of instruments and joints in such a test rig, it is obvious that in a test rig of this size (16-in-diam test section) special sealing methods must be considered to keep the leakage small.

To operate the test loop at 97 per cent (by weight) pure Freon it was necessary to use a purification cycle. Gas mixture was drawn off to the Freon compressor after which it was condensed and the noncondensables (mostly entrained air and vapors) were dumped to the atmosphere. Freon liquid was supplied to the test loop. The purification cycle also was used to control the quantity of fluid in the test loop in the following manner: (i) An expansion-type valve was set to expand the incoming liquid to a gas; (ii) an automatic throttling valve, sensitive to the inlet-tank pressure, controlled the quantity of fluid being taken out of the test rig into the purification cycle. The expansion valve could be adjusted to keep the flow quantity within the control range of the valve in the suction line.

Temperature level in the entrance tank was controlled automatically by the flow of water to the coolers. Very good control of the test-loop conditions was possible with this automatic system. It should be pointed out that with this system the quantity of gas in the test loop must be changed for each test point. The variation was kept as small as possible by the small volume of high-pressure gases between the compressor rotor and the throttle valve. This system responded to these changes very rapidly and no undue delay in setting test points was involved.

Freon purity was measured continuously and indicated on a chart recorder which was calibrated to read percentage of Freon. A sampling line in the entrance tank carried a small quantity of test gas to the analyzer, which compared the electrical conductivity of this gas with that of pure Freon. This analyzer was very sensitive and showed considerable advantage over other devices in use. The use of this device greatly facilitated the tests.

One other problem that was encountered seems important. A certain amount of oil leakage into the test gas is inevitable. In a closed test loop this contamination may build up until it cannot be tolerated. In the test rig described, the trouble occurred in the purification circuit. The oil vapor carried to the Freon system formed a wax at the condensation temperature of Freon-12. In this case a refrigeration-system oil was substituted in the test rig. The oil vapor thus removed in the purification cycle was sufficient to keep the test loop reasonably clear of oil.

From these experimental problems some implications can be drawn concerning the problems of testing with gases other than air which have been blended to obtain special values of specific-heat ratios. It seems apparent that a purification system is necessary to charge the system and maintain the necessary concentration and control as well as salvaging the gas mixtures at the end of a test run. The present concept of a purification cycle is to separate the gases and remix them in their proper proportions. Some gases may be considerably more difficult to separate than the Freon from the leakage air encountered in this test loop. It is necessary to know with considerable accuracy that the gas in the test loop is of the proper mixture. Thus a good method of analyzing these gas mixtures must be worked out. It would seem that chemical-absorption or density-balance methods are too time consuming and tedious. It should not be necessary to operate a compressor test facility for extended periods while the test gas is analyzed and corrected. Thus there are some problems involved which indicate that testing in such gas mixtures would require considerable development, time, and money. It does appear that the use of special gas mixtures for testing compressors, under certain circumstances, may be economical over a very long period. It certainly cannot be considered as a quick fix for extending test facilities.

Conclusion

The following concluding remarks can be made concerning the translation of compressor performance from one gas to another and the use of special test mediums for obtaining compressor performance:

1 The methods of translating compressor performance from one gas to another are rather long and cumbersome. However, if approximate results are desired, they can be handled quite simply.

2 As the number of blade rows increases, the translation of compressor performance becomes much more difficult because geometrically similar flow cannot be set at all points.

3 Fundamental data obtained in Freon or other test mediums may be useful in designing compressors for air. However, the substitution of these test mediums to extend the range of a test facility may introduce additional problems.

4 Gas mixtures with a specific-heat ratio equivalent to the design gas may be used as the test medium and would avoid the problem of translating the compressor performance. The mechanical problems involved in a test stand using such a test gas and the problems involved in obtaining the gas mixture and maintaining it would probably tend to decrease the usefulness of such a test medium.

BIBLIOGRAPHY

1 "Investigation of an Experimental Supersonic Axial-Flow Compressor," by J. R. Erwin and L. C. Wright, NACA RM L6J010, 1946.

2 "Fluid Mechanics of Turbomachinery," by G. F. Wislicenus, McGraw-Hill Book Company, Inc., New York, N. Y., and London, England, 1947.

3 "Use of Freon-12 as a Fluid for Aerodynamic Testing," by P. W. Huber, NACA TM 1024, 1947.

4 "Performance of 24-Inch Supersonic Axial-Flow Compressor in Air. I—Performance of Compressor Rotor at Design Speed of 1600 Feet per Second," by W. K. Ritter and I. A. Johnsen, NACA RM E7L10, 1948.

5 "Testing of Multistage Compressors With Heavy Gas," by F. S. Weing, AAF, Air Materiel Command, Wright Field Technical Report F-TR-2143-ND, June, 1947.

6 "Some Possibilities of Using Gas Mixtures Other Than Air in Aerodynamic Research," by D. R. Chapman, NACA-AAL-TN 3226, August, 1954.

Appendix

Equations and Methods Necessary for Translation of Compressor Performance

The prediction of compressor performance in some gas other than the design gas depends on matching certain conditions between the two gases. Quite obviously geometrically similar flow at the blade-entrance section should be matched as closely as possible and the variations in flow through the rotor passage computed. Geometrically similar flow at the blade-row entrance would require that the relative flow Mach number and angle be the same in the two gases (as illustrated in Fig. 2).

The first point to be determined is the rotational speed at which the inlet equivalent Mach diagrams are obtained. The relative Mach number and flow angle must be the same and thus

$$\left(\frac{\omega r_1}{a_1}\right)_{dg} = \left(\frac{\omega r_1}{a_1}\right)_{tm}$$

$$\frac{(\omega r_1)_{tm}}{(\omega r_1)_{dg}} = \frac{(a_1)_{tm}}{(a_1)_{dg}} = \frac{(a_1/a_0)_{tm} a_{0,tm}}{\left(\frac{a_1}{a_0}\right)_{dg} a_{0,dg}}$$

which when considered as a perfect gas

$$\frac{(\omega r_1)_{tm}}{(\omega r_1)_{dg}} = \sqrt{\left(\frac{\gamma_{tm} R_{tm} T_{tm}}{\gamma_{dg} R_{dg} T_{dg}}\right)} \left(\frac{1 + \frac{\gamma_{dg}-1}{2} M_{1,dg}^2}{1 + \frac{\gamma_{tm}-1}{2} M_{1,tm}^2}\right)^{1/2}$$

and reducing this to equivalent speed

$$\frac{(N/\sqrt{\theta})_{tm}}{(N/\sqrt{\theta})_{dg}} = \left(\frac{\gamma_{tm} R_{tm}}{\gamma_{dg} R_{dg}}\right)^{1/2} \left(\frac{1 + \frac{\gamma_{dg}-1}{2} M_{1,dg}^2}{1 + \frac{\gamma_{tm}-1}{1} M_{1,tm}^2}\right)^{1/2} \quad [6]$$

Since the equivalent inlet Mach diagrams are assumed we can now determine the relation between the gas flow rates of the two gases by writing

$$\frac{W_{tm}}{W_{dg}} = \frac{(\rho V \cos \beta_1 A)_{tm}}{(\rho V \cos \beta_1 A)_{dg}} = \frac{(\rho M_1 \cos \beta_1 a_1 A)_{tm}}{(\rho M_1 \cos \beta_1 a_1 A)_{dg}}$$

since $M_{1,tm} = M_{1,dg}$ and $\beta_1 = \beta_1$ and $A = A$

$$= \frac{\rho_{tm} a_{1,tm}}{\rho_{dg} a_{1,dg}} = \frac{(\rho a)_{tm}}{(\rho a)_{dg}}$$

$$= \frac{\left[\left(\frac{\rho}{\rho_0}\right)\left(\frac{a_1}{a_0}\right)\left(\frac{P}{RT}\right)\sqrt{(\gamma RT)}\right]_{tm}}{\left[\left(\frac{\rho}{\rho_0}\right)\left(\frac{a_1}{a_0}\right)\left(\frac{P}{RT_1}\right)\sqrt{(\gamma RT)}\right]_{dg}}$$

On reducing this to standard sea-level weight flow and substituting

$$\frac{\left(\frac{W\sqrt{\theta}}{\delta}\right)_{tm}}{\left(\frac{W\sqrt{\theta}}{\delta}\right)_{dg}} = \sqrt{\frac{(\gamma/R)_{tm}}{(\gamma/R)_{dg}}} \frac{\left[\left(1 + \frac{\gamma-1}{2} M_1^2\right)^{\frac{1+\gamma}{2(\gamma-1)}}\right]_{dg}}{\left[\left(1 + \frac{\gamma-1}{2} M_1^2\right)^{\frac{1+\gamma}{2(\gamma-1)}}\right]_{tm}} \quad [7]$$

Having set operating conditions to obtain equivalent flow at the inlet, the flow through the rotor must be considered. The continuity equation can be written for the rotating stream tube shown in Fig. 5 for a given gas

$$\rho_1' M_1' a_1' A_1' = \rho_2' M_2' a_2' A_2'$$

(Since stream conditions are equivalent in the relative and absolute reference frames the primes can be dropped from a and ρ but not from a_0 and ρ_0)

$$\frac{M_1'}{M_2'} = \frac{\rho_2 a_2 A_2'}{\rho_1 a_1 A_1'} = \frac{a_{2,0}' \left(\frac{a_2}{a_{2,0}'}\right) \rho_{2,0}' \left(\frac{\rho_2}{\rho_{2,0}'}\right) A_2'}{a_{1,0}' \left(\frac{a_1}{a_{1,0}'}\right) \rho_{1,0}' \left(\frac{\rho_1}{\rho_{1,0}'}\right) A_1'}$$

then substituting for a_0', ρ_0', the perfect gas law and the isentropic relation for the ratios of velocities of sound and density as a function of Mach number

$$\frac{M_1'}{M_2'} = \sqrt{\left(\frac{T_1'}{T_2'}\right)} \frac{P_2'}{P_1'} \frac{A_2'}{A_1'} \left(\frac{1 + \frac{\gamma-1}{2} M_2'^2}{1 + \frac{\gamma-1}{2} M_1'^2}\right)^{-\frac{\gamma+1}{2(\gamma-1)}}$$

Now if we write this for the two gases, presuming that we have set $(M_1')_{tm} = (M_1')_{dg}$, and that the effective flow areas are equal $(A_2'/A)_{tm} = (A_2/A_1')_{dg}$

$$\frac{M_{2,tm}'}{M_{2,dg}'} = \sqrt{\left[\frac{(T_1'/T_2')_{dg}}{(T_1'/T_2')_{tm}}\right]} \frac{(P_2'/P_1')_{dg}}{(P_2'/P_1')_{tm}} \frac{\left[\left(\frac{1 + \frac{\gamma-1}{2} M_2'^2}{1 + \frac{\gamma-1}{2} M_1'^2}\right)^{-\frac{\gamma+1}{2(\gamma-1)}}\right]_{dg}}{\left[\left(\frac{1 + \frac{\gamma-1}{2} M_2'^2}{1 + \frac{\gamma-1}{2} M_1'^2}\right)^{-\frac{\gamma+1}{2(\gamma-1)}}\right]_{tm}}$$

Now collecting the unknown terms including relative discharge Mach number in the test medium

$$M_{2,tm}' \left[\left(1 + \frac{\gamma-1}{2} M_2'^2\right)^{-\frac{\gamma+1}{2(\gamma-1)}}\right]_{tm} = \sqrt{\left[\frac{(T_1'/T_2')_{dg}}{(T_1'/T_2')_{tm}}\right]\left[\frac{(P_2'/P_1')_{dg}}{(P_2'/P_1')_{tm}}\right]} \left[\left(1 + \frac{\gamma-1}{2} M_2'^2\right)^{-\frac{\gamma+1}{2(\gamma-1)}}\right]_{dg}$$

$$M_{2,dg}' \frac{\left\{\left[\left(1 + \frac{\gamma-1}{2} M_2'^2\right)\right]^{-\frac{\gamma+1}{2(\gamma-1)}}\right\}_{dg}}{\left[\frac{1}{\left(1 + \frac{\gamma-1}{2} M_1'^2\right)^{-\frac{\gamma+1}{2(\gamma-1)}}}\right]_{tm}} \quad [8]$$

Thus the terms on the right-hand side of the equation are known from the design gas and the conditions set in the test medium except for the ratios of relative total temperatures and pressures. Without heat transfer the relative total temperature ratio can be obtained from the known inlet and outlet radius, rotational speed, and gas properties, as indicated in the following equation

$$T_2'/T_1' = 1 + \left(\frac{\omega r_2}{a_{0,1}'}\right)^2 \frac{\gamma-1}{2} \left[1 - \left(\frac{r_1}{r_2}\right)^2\right]$$

and an ideal relative total pressure ratio

$$(P_2'/P_1')_{ideal} = \left(\frac{T_2'}{T_1'}\right)^{\frac{\gamma}{\gamma-1}}$$

$$= \left\{1 + \left(\frac{\omega r_2}{a_{0,1}'}\right)^2 \frac{\gamma-1}{2} \left[1 - \left(\frac{r_1}{r_2}\right)^2\right]\right\}^{\frac{\gamma}{\gamma-1}}$$

A loss in total pressure ratio must be considered

$$P_2'/P_1' = \left(\frac{P_2'}{P_1'}\right)_{ideal} \left(\frac{P_2'}{P_{2,ideal}'}\right)$$

If the loss in total pressure is considered as a fractional value of the inlet dynamic pressure and assumed equal for the two cases

$$\left(\frac{P_{2,ideal}' - P_2'}{P_1' - p_1'}\right)_{tm} = \left(\frac{P_{2,ideal}' - P_2'}{P_1' - p_1'}\right)_{dg}$$

The total pressure loss may be assigned in this manner to a design or it may be determined from test data in one of the gases. Sufficient information is then available to determine the relative discharge Mach number in the test medium.

Since the relative discharge Mach number and total tempera-

ture are known the relative stream velocity of sound can be determined (also being equal to the absolute stream velocity of sound). Thus the discharge relative velocity is known and, assuming that the relative flow angle is the same in the two gases, these can be combined with the discharge rotational speed to compute the absolute discharge vector diagram.

The following determination of performance values is then possible in either gas. The total pressure ratio could be obtained by a series of pressure ratios for the required Mach numbers and the loss and relative total pressure changes as shown in the following

$$\frac{P_2}{P_1} = \frac{P_2}{p_2} \times \frac{p_2}{P_2'} \times \frac{P_2'}{P_{2,\text{ideal}}'} \times \frac{P_{2,\text{ideal}}'}{P_1'} \times \frac{P_1'}{p_1'} \times \frac{p_1}{P_1}$$

where the relative and absolute stream pressures are equal at inlet and outlet and for temperature ratio

$$\frac{T_2}{T_1} = \frac{T_2}{t_2} \times \frac{t_2'}{T_2'} \times \frac{T_2'}{T_1'} \times \frac{T_1'}{t_1} \times \frac{t_1}{T_1}$$

In the foregoing equations all ratios of pressure and temperature are determined as a function of the Mach number or otherwise previously determined. The adiabatic efficiency thus obtained in both gases can be determined

$$\eta_{\text{ad}} = \frac{(P_2/P_1)^{\frac{\gamma}{\gamma-1}} - 1}{T_2/T_1 - 1}$$

and the ratio of power requirement to operate at geometrically similar flow conditions can be determined from

$$\frac{(\text{HP})_{\text{tm}}}{(\text{HP})_{\text{dg}}} = \frac{\left[\left(\frac{T_2}{T_1} - 1\right) T_1 \left(\frac{\gamma R}{\gamma - 1}\right) \frac{W\sqrt{\theta}}{\delta}\right]_{\text{tm}}}{\left[\left(\frac{T_2}{T_1} - 1\right) T_1 \left(\frac{\gamma R}{\gamma - 1}\right) \frac{W\sqrt{\theta}}{\delta}\right]_{\text{dg}}}$$

$$\frac{\left(\frac{\text{HP}}{\theta}\right)_{\text{tm}}}{\left(\frac{\text{HP}}{\theta}\right)_{\text{dg}}} = \frac{\left[\left(\frac{\gamma R}{\gamma - 1}\right)\left(\frac{T_2}{T_1} - 1\right)\right]_{\text{tm}}}{\left[\left(\frac{\gamma R}{\gamma - 1}\right)\left(\frac{T_2}{T_1} - 1\right)\right]_{\text{dg}}}$$

The performance values and flow condition in the two gases can be compared.

Discussion

J. T. Hamrick.[4] The authors are to be congratulated on the presentation of an excellent paper. It points up many of the problems of using simulated gases and should serve as a stimulus to manufacturers now faced with the problem of testing with simulated gases, and also to users faced with the problem of pumping gases which vary in ratio of specific heats over a period of time. For such cases, extreme care should be used in the aerodynamic design of the compressor, and especially the impeller, in order to assure as wide a range of operation as possible. For example, where surge is caused by an unstable eddy on the driving face of a centrifugal impeller (see Fig. 2 of paper by Hamrick and Beede[5]), a change in the gas properties may result in a significant shift in the compressor range. This effect was noted[5] with the injection of water into the compressor inlet. Injection of water did not significantly change the inlet conditions in the impeller, but the greater compression within the impeller with the injection of water led to a lower volume flow rate through the latter part of the impeller passage and the occurrence of the unstable eddy at a higher weight flow than previously. This trouble was corrected by the addition of splitter vanes to reduce the aerodynamic loading and eliminate the eddy. Even after this correction, the surge line moved to a higher weight flow than previously because of the change of flow angle into the diffuser vanes, but the shift in this case could be tolerated.

For multistage compressors, variation in the gas properties could result in rotating stall in the latter stages, a phenomenon which has led to embarrassing blade failures in jet engines. In fact, variation in the specific-heat ratio of the test gas might be an excellent method of inducing rotating stall in a multistage compressor for research studies or it could be used by the manufacturer to indicate the proximity of a rotating-stall condition in a compressor.

A further use of the variation in the specific-heat ratio is the simulation of boundary-layer effects in a compressor. For example, if in either an axial-flow or centrifugal compressor, a variation in specific-heat ratio resulted in a maximum efficiency at a ratio other than the design specific-heat ratio, then an improper allowance for boundary layer has been made.

---

[4] Engineering Specialist, Thompson Products Inc., Cleveland, Ohio.

[5] "Some Investigation With Wet Compression," by J. T. Hamrick and W. L. Beede, Trans. ASME, vol. 75, 1953, pp. 409–420.

Heat-Transfer Rates to Crossflowing Mercury in a Staggered Tube Bank[1]—I

By R. J. HOE,[2] D. DROPKIN,[3] and O. E. DWYER[4]

An experimental heat-transfer program has been under way for some time at the Brookhaven National Laboratory, the general purpose of which is to obtain and correlate heat-transfer coefficients for liquid metals flowing outside of tube banks. This paper covers one phase of the program; i.e., the measurement of heat-transfer coefficients for mercury flowing normal to a staggered tube bank. The factors studied were (*a*) linear velocity, (*b*) tube location in the bank, (*c*) circumferential variation of the local coefficient for a single tube, (*d*) type of contact, i.e., "wetting" versus "nonwetting," and (*e*) pressure drop.

Nomenclature

The following nomenclature is used in the paper:

A = heat-transfer area per tube, sq ft
A_F = cross-sectional area for flow, sq ft
A_H = heat-transfer area, sq ft
C = a constant = $3.413\,IE/k_c$ (inside surface area of tube)
C_p = specific heat of fluid evaluated at average bulk temperature, Btu/lb-deg F
C_{pf} = specific heat of fluid evaluated at average film temperature, Btu/lb-deg F
D_o = outside diameter of tube, ft
De' = equivalent diameter = 4 (free volume)/(surface area of tubes), ft
E = voltage drop in heating element corresponding to area A
f = friction factor, dimensionless, defined by Equation [14]
g_0 = conversion factor, 4.17×10^8 (lb mass) (ft)/(hr)^2 (lb force)
h = local heat-transfer coefficient, Btu/hr-sq ft-deg F
$\bar{h}$ = average heat-transfer coefficient for a given tube, Btu/hr-sq ft-deg F
I = electric current, amp
k = thermal conductivity of fluid evaluated at average bulk temperature, Btu/hr-ft-deg F
k_c = thermal conductivity of copper, Btu/hr-ft deg F
k_f = thermal conductivity of fluid evaluated at average film temperature, Btu/hr-ft deg F
N = number of rows of tubes across which fluid flows
N_{Pr} = Prandtl number, $C_{pf}\mu_f/k_f$
N_{Re} = Reynolds number, $D_oV_{max}\rho_f/\mu_f$
Δp = pressure drop, psf
P = tube pitch, distance between centers, ft
q = $3.413\,IE$ = rate of heat transfer from whole tube, Btu/hr
r = radius of tube, ft
r_i = inside radius of tube, ft
r_o = outside radius of tube, ft
t_b = bulk temperature of fluid, deg F
t_w = local outside wall temperature of tube, deg F
T = tube-wall temperature at radius r, deg F
V = shell-side velocity of fluid flowing parallel to tubes, fph
V_s = geometric mean of the velocity across tubes and that around baffles, fph
V_{max} = shell-side velocity of fluid across tube bank and based on minimum flow area, fph
ρ = density of fluid, pcf
μ = absolute viscosity of fluid evaluated at average bulk temperature, lb/(hr)(ft)
μ_f = absolute viscosity of fluid evaluated at average film temperature, lb/(hr)(ft)
θ = angle from forward stagnation point, deg

Introduction

The use of liquid metals as heat-transfer media for nuclear reactors has received considerable attention in recent years. Liquid metals have the following important advantages over other types of coolants: High reactor temperatures are possible at low pressures, excellent heat-transfer characteristics, and absence of radiation damage. At the present time, there are several reactor types (either operating, under construction, or in the design stage) which involve the use of liquid metals—both alkali and heavy.

The theory and practice of liquid-metal heat transfer are not as yet on such firm bases as those for conventional fluids, and consequently heat-exchange equipment cannot be designed with the desired degree of precision. A considerable amount of liquid-metal heat-transfer research has been carried out in recent years; but nearly all of it has been confined to flow inside tubes. However, crossflow information is also necessary for the design of certain types of heat-exchanger equipment, particularly shell-and-tube heat exchangers.

The present investigation was undertaken for the purpose of measuring crossflow heat-transfer coefficients for mercury flowing normal to a staggered tube bank. It was assumed that data obtained on mercury would be applicable to the other heavy metals, owing to the fact that they are all quite similar in so far as their thermal and physical properties are concerned. For experimental work, mercury has the great advantage over the other heavy metals and their alloys, of being liquid at room temperature. This greatly simplified research equipment and experimental methods.

The experimental conditions and scope of the present investigation are summarized briefly in Table 1. Owing to the fact that the bulk temperature of the mercury could not be varied appreciably in the present study, Prandtl number was not studied as an independent variable.

---

[1] This paper is based upon a research project conducted at the Brookhaven National Laboratory by R. J. Hoe for the PhD degree from Cornell University. The work was sponsored by the U. S. Atomic Energy Commission.

[2] Engineer, Knolls Atomic Power Laboratory, General Electric Company, Schenectady, N. Y.

[3] Associate Professor of Mechanical Engineering, Department of Thermal Engineering, Cornell University, Ithaca, N. Y.

[4] Head, Chemical Engineering Division, Nuclear Engineering Department, Brookhaven National Laboratory, Upton, N. Y.

Contributed by the Heat Transfer Division and presented at the Semi-Annual Meeting, Cleveland, Ohio, June 17–21, 1956, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, March 1, 1956. Paper No. 56—SA-28.

TABLE 1 EXPERIMENTAL CONDITIONS AND SCOPE OF INVESTIGATION

Tube size, OD, in.	0.500
Length of tubes, vertical, in.	7.0 (approx)
Tube spacing	Equilateral triangle
Tube pitch, in.	11/16
Pitch/diameter	1.375
Tube bank size	7 tubes wide × 10 tubes deep
Bulk mercury temperature, deg F	100 (approx)
Maximum temperature rise of mercury in passing through tube bank, deg F	1/2 (approx)
Prandtl number	0.022
Reynolds number range	15,000 to 83,000
Linear-velocity range, V_{max}, fps	0.42 to 2.25

Review

The mechanism of heat transfer in a turbulently flowing liquid-metal stream is quite different from that in a nonmetallic fluid. The principal difference is due to the great disparity in the thermal conductivities of the two types of fluids. Whereas in the case of ordinary fluids, the majority of the heat is transferred to the turbulent stream by eddy conductivity, in the case of liquid metals it is by molecular and ionic conductivity. For this reason, the relationships which have been developed for nonmetals will not apply to liquid metals.

Little progress has been made in determining crossflow coefficients for tube banks by theoretical analysis, owing to the complexity of the flow patterns. However, in the case of ordinary fluids, considerable success has been achieved with the use of dimensional analysis.

Prior to this study, to the authors' knowledge no heat-transfer data had been obtained on a liquid metal flowing normal to a tube bank. The two closest systems, on which results have been reported are (*a*) an unbaffled shell-and-tube heat exchanger with NaK on the shell side, and (*b*) a half-moon-baffled shell-and-tube heat exchanger also with NaK on the shell side. In the first case, Brooks and Rosenblatt (1)[5] correlated shell-side coefficients by the equation

$$\frac{hD_o}{k} = 61.2\left[\left(\frac{D_oV\rho C_p}{k}\right)\left(\frac{A_F}{A_B}\right)^2\right]^{0.60} \quad [1]$$

with a mean deviation of ±15 per cent for the data; and in the second case, Tidball (2) correlated shell-side coefficients by Donohue's (3) equation, the corresponding average deviation being ±25 per cent. Donohue's equation is

$$\frac{hD_o}{k} = 0.19(De')^{0.6}\left(\frac{D_oV_o\rho}{\mu}\right)^{0.6}\left(\frac{C_p\mu}{k}\right)^{1/3} \quad [2]$$

and was derived from data on water and oil flowing in baffled shell-and-tube heat exchangers. Brooks and Rosenblatt, Tidball, and Donohue all measured over-all heat-transfer coefficients and subtracted the resistances of the tube wall and tube-side fluid from the measured over-all resistances to calculate the shell-side coefficients.

Theoretical Considerations

The local heat-transfer coefficient h is defined by the equation

$$h = \frac{dq}{(t_w - t_b)dA} \quad [3]$$

If it is assumed that circumferential heat flow is relatively small, Equation [3] can be written

$$h = \frac{q}{(t_w - t_b)A} \quad [3a]$$

[5] Numbers in parentheses refer to the Bibliography at the end of the paper.

The average coefficient $\bar{h}$ for a given tube can be defined in two ways: By the equation

$$\bar{h} = \frac{\int_0^{360} h\,d\theta}{360} \quad [4]$$

and by the equation

$$\bar{h} = \frac{360q}{A\int_0^{360}(t_w - t_b)d\theta} \quad [5]$$

In this study $\bar{h}$ was evaluated by Equation [4] and h by Equation [3*a*].

Combining Equations [3*a*] and [4] we get

$$\bar{h} = \frac{\int_0^{360}[q/A(t_w - t_b)]\,d\theta}{360} \quad [6]$$

Now, comparing Equations [5] and [6], it is seen that $(t_w - t_b)$ is being averaged in the former, while $1/(t_w - t_b)$ is being averaged in the latter; otherwise, the two equations are the same. The values of $\bar{h}$ from these two equations will differ depending upon the variation of t_w with θ; the greater the variation the greater the difference. In the present study, the variation of t_w with θ was such that it is estimated that $\bar{h}$ by Equation [6] is, in general, about 15 to 20 per cent above $\bar{h}$ by Equation [5]. In this paper, values of $\bar{h}$ were calculated by means of Equation [6].

The method of calculating local coefficients from the experimental data when circumferential heat flow is taken into account is outlined in the following:

For steady state and no heat transfer along the axis of the tube, the partial differential equation in cylindrical co-ordinates for two-dimensional heat flow is

$$\frac{\partial^2 T}{\partial r^2} + \frac{1}{r}\frac{\partial T}{\partial r} + \frac{1}{r^2}\frac{\partial^2 T}{\partial \theta^2} = 0 \quad [7]$$

Assumed boundary conditions are

$$\left(\frac{\partial T}{\partial r}\right)_{r=r_i} = C$$

and

$$T_{(r_o,\theta)} = a_0 + \sum_{n=1}^{\infty} a_n \cos n\theta$$

Integration of Equation [7] yields

$$T_{(r,\theta)} = A_0 + B_0 \ln r + \sum_{n=1}^{\infty}(A_n r^n + B_n r^{-n})\cos n\theta \quad [8]$$

Using the boundary conditions to determine the coefficients and constants, Equation [8] becomes

$$T_{(r,\theta)} = a_0 + Cr_i \ln\frac{r}{r_o} + \sum_{n=1}^{\infty} a_n \frac{r_o^n}{r^n}\left(\frac{r^{2n} + r_i^{2n}}{r_o^{2n} + r_i^{2n}}\right)\cos n\theta \quad [9]$$

Differentiating Equation [9] with respect to r and substituting r_o for r, we get

$$\left(\frac{\partial T}{\partial r}\right)_{r=r_o} = \frac{Cr_i}{r_o} + \sum_{n=1}^{\infty}\frac{na_n}{r_o}\left(\frac{r_o^{2n} - r_i^{2n}}{r_o^{2n} + r_i^{2n}}\right)\cos n\theta \quad [10]$$

and finally, at a particular value of θ

$$\left(\frac{dq}{dA}\right)_{r=r_0} = -k_c\left(\frac{\partial T}{\partial r}\right)_{r=r_0} = h(t_w - t_b) \quad [11]$$

Equation [11] gives the true local heat-transfer rate, as it takes into account both radial and circumferential heat flow through the tube wall.

Experimental Equipment

The experimental equipment used in the present investigation consisted of a liquid-metal circulation loop with the following major components:

1 Pipe loop—1 1/2-in. schedule-80 alloy-steel pipe (5 per cent Cr, 1/2 per cent Mo) connections are flanged with tapered-insert steel gaskets.

2 Pump—canned motor molten metal pump—direct-drive centrifugal 15 hp 220 volt a-c.

3 Test section—steel vessel with stainless-steel insert consisting of straightening tubes of type 304 SS 1 in. × 8 in. long; a calming section; and the tube bundle in a rectangular lattice.

4 Cooling section—double-pipe heat exchanger consisting of 2 1/2-in. schedule-40 steel pipe 5 ft long over a section of the primary loop. Site cooling water was used as coolant.

5 Motor generator set—7 1/2 hp 440 volt a-c motor with 5-kw, 240-volt d-c generator. This set is used to supply direct current to the heating element.

A flow diagram of the loop is shown in Fig. 1. The tube bank, Fig. 2, consisted of 70 tubes, 1/2-in. OD with an active length of 7 in. The lattice was an equilateral triangular arrangement, seven rows wide by ten rows deep, with a pitch or distance between centers of 11/16 in. Twelve of the tubes were available for test, as indicated in Fig. 2, while the remainder, made of 1/2-in. type 347 stainless steel, were dummies.

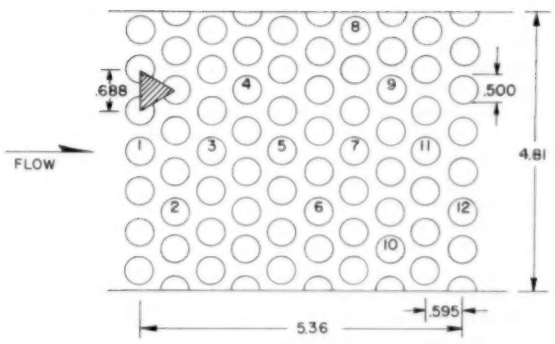


Fig. 2 Plan View of Tube Bank. All Dimensions Are in Inches

The test elements consisted of 1/2-in-OD, type K copper tubes, 0.049 in. thick and 10 1/2 in. long. Most of the elements were chrome-plated on the outside wall with a 0.0005-in. chrome plate to prevent amalgamation of the mercury and the copper. However, some elements were left bare so that the effect of "wetting" could be determined. The tubes were sealed in the lattice by means of packing glands with Teflon gaskets and packing nuts.

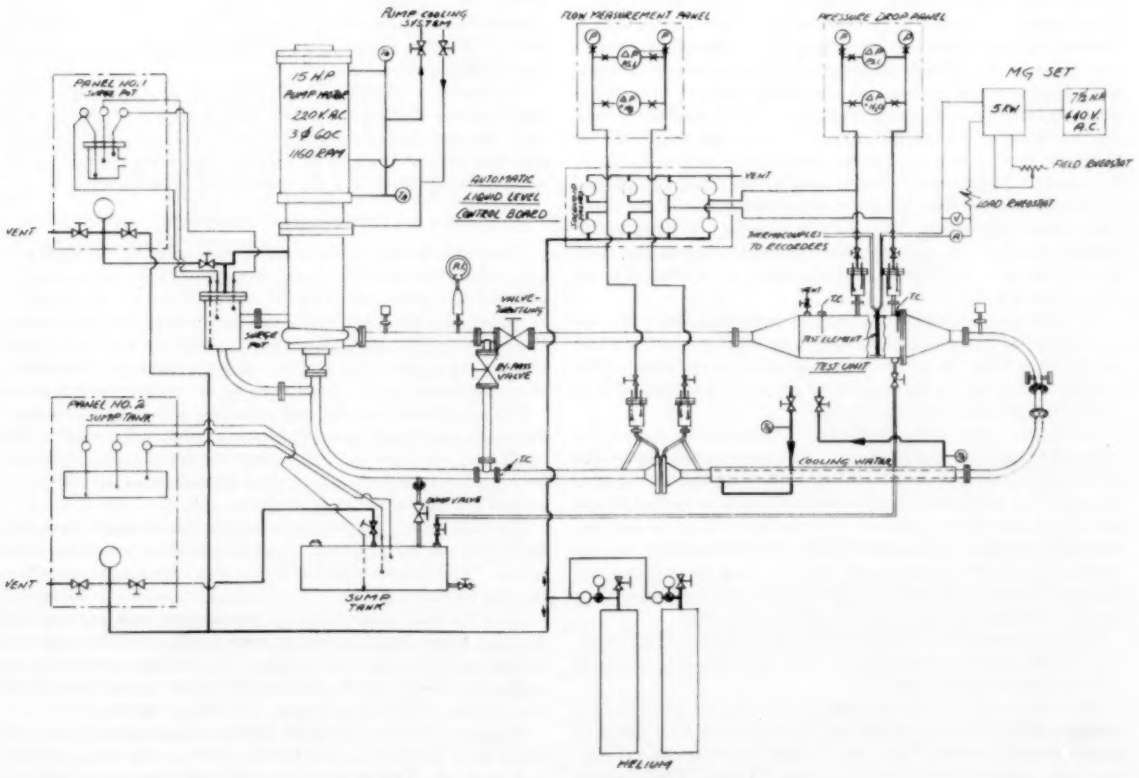


Fig. 1 Schematic Diagram of Test Equipment

The surface temperature of the tube was measured by means of thermocouples imbedded in the walls. The method of making these thermocouple tubes was as follows: First, a fine slit $6^1/_2$ in. long × 0.006 in. wide × 0.015 in. deep was milled in the outside wall of the tube. Into this slit was placed a no. 36 or no. 40 gage constantan thermocouple wire which had been specially coated with "Formvar" insulation. A bare copper wire of a diameter slightly larger than the slot width was then forced into the slot. An ohmmeter was attached to the constantan wire and the tube to detect a break in the insulation while the copper was being worked into the slot. If a break occurred, the wire was replaced. When the copper was properly wedged into the slot, it was then worked into the surface at the point where the constantan wire came out of the slot. This process stripped the insulation off the constantan wire and caused a metallic bond between the copper wire, the constantan wire, and the tube wall immediately at the surface. A thermocouple junction was thereby formed near the tube-wall surface at the middle of the tube section exposed to the crossflow of mercury. In all, nine thermocouples, 40 deg apart, were placed around the circumference of the tube. The fine slots used for the thermocouple would have very little effect upon the heat-flux pattern and an accurate determination of the surface temperature could be obtained. The tube was then polished and chrome-plated or, in a few cases, left unplated.

The temperatures were measured and recorded on high-speed recorders so that fluctuations in the temperatures could be observed readily. All thermocouples and recorders were calibrated.

The heat source consisted of an electrical heater inserted inside the test element, electrical energy being supplied from the 5-kw, 240-volt, d-c generator. Because of the high heat flux required and the small space available, special heaters had to be developed. These heaters consisted of corrugated molybdenum wire wound evenly on a 6-mm-diam capillary Vycor tube. An outer tube of 9-mm OD and 7-mm ID was then slipped over the assembly, heated, and collapsed under vacuum. It was found necessary to flush the assembly with an inert gas several times before sealing, to prevent oxidation of the molybdenum wire at the high temperatures. Platinum wires were welded to the molybdenum before sealing in the quartz. Nickel wire was then used for the leads. To further insure against oxidation, helium was supplied as a cover gas to the heater during operation. The electrical heater was fitted inside the test element with mercury added to the annulus to give good thermal contact. A diagram of the assembly of the test element is shown in Fig. 3.

The bulk temperature of the mercury was measured with copper-constantan thermocouples which were located upstream and downstream from the lattice and connected to recorders. The mercury flow rate was determined by means of a calibrated flow nozzle and differential pressure gages.

The power input to the heater was determined by means of a calibrated voltmeter and ammeter. The pressure drop across the tube bank was measured with precision differential-pressure gages. In measuring pressures, the mercury levels were controlled in seal pots by means of two probes. The probes in a given seal pot operated through relays and solenoid valves to regulate the gas pressure above the mercury and thereby bring the level of the mercury between the probes. Differential-pressure gages then read the difference in gas pressure in the two seal pots.

Auxiliaries included ventilation equipment for removal of mercury vapor, control equipment for the M-G set, and mercury-vapor detection equipment.

Since heat-transfer experiments may be affected adversely by foreign materials in the system, great precautions were taken to obtain a clean system. The inside of the loop was cleaned consecutively with steel wool and wire brush, "Turco," a commercial acid rust-removing solution, hot water and detergent, water, trichloroethylene, and ethyl alcohol. After circulating the water, the system was dried and evacuated. The same was done after the alcohol wash. Then, while under vacuum, the system was heated with an acetylene torch to remove all the alcohol. The vacuum-heating process was resumed after filling the system with helium. Finally, the system was filled with helium to a pressure slightly above atmospheric.

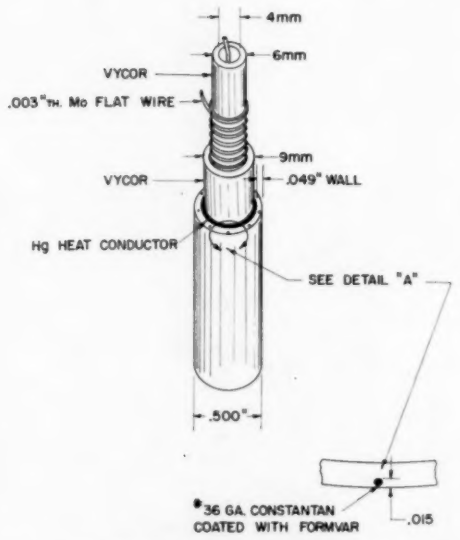


FIG. 3 CUTAWAY DRAWING SHOWING CONSTRUCTION AND ASSEMBLY OF TEST ELEMENT

The mercury, which was triple distilled, was pumped into the sump under a blanket of helium. About 1400 lb of mercury were used, this amount being enough to insure that the intake to the fill pipe was covered at all times, thereby preventing floating foreign material being raised into the test section from the sump tank.

EXPERIMENTAL PROCEDURE

The first series of tests consisted of measuring the pressure drop across the lattice as a function of flow rate. During these tests no heat-transfer measurements were made. The flow was regulated by either throttling or bypassing the flow to the test section. Pressure-drop data were taken over the 15,000 to 80,000 Reynolds-number range, the lower limit being determined by the precision of measurement and the upper one by the capacity of the pump.

The heat-transfer coefficients were determined by supplying a measured heat input to a test element; and, when steady-state conditions were reached, measuring the temperature difference between the tube surface and the bulk stream at several points around the circumference of the tube.

The bulk temperature of the mercury was obtained by taking the average of the temperatures upstream and downstream of the lattice. This temperature difference was very small—usually in the order of 0.5 deg.

Since the bulk temperature of the mercury was near ambient, the heat losses from the system were small, thus allowing equilibrium to be reached quite rapidly. The fluid properties used for evaluating the Reynolds number were taken at the average film temperature, which was in general between 105 to 115 F.

Data were obtained using the chrome-plated element for 11 different tube locations in the lattice, several runs being made at each location. Two runs were made using a bare copper element at a central tube location and one run at the No. 1 tube location

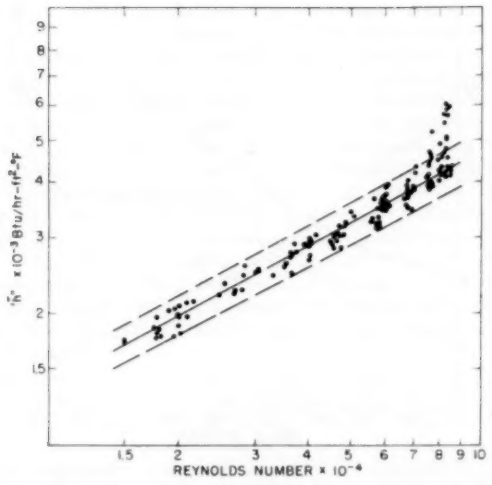


FIG. 4 AVERAGE HEAT-TRANSFER COEFFICIENTS FOR CHROME-PLATED TUBES IN INTERIOR OF TUBE BANK (LOCATIONS 3, 4, 5, 6, AND 7)

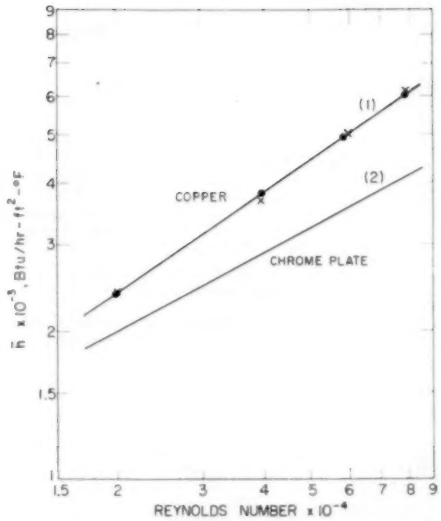


FIG. 5 HEAT-TRANSFER COEFFICIENTS UNDER WETTING AND NONWETTING CONDITIONS IN INTERIOR OF TUBE BANK

which was at the front of the bank. Since mercury amalgamates readily with copper, these tests were destructive. A copper tube lasted about two days before the thermocouples malfunctioned.

RESULTS

Interior of Tube Bank. The values of the physical properties used in the correlation work discussed later were taken from the "Liquid Metals Handbook" (4) and evaluated at average film temperature, which was in general between 105 and 115 F. These values, in the present case, are not appreciably different from those evaluated at the average bulk temperature.

Fig. 4 shows the observed average coefficient on a single-tube basis $\bar{h}$ for tube locations 3, 4, 5, 6, 7 (see Fig. 2) plotted against Reynolds number. Within experimental error all locations gave the same results. The straight line drawn by sight through the points is represented by the equation

$$\bar{h} = 11.6(D_o V_{max}\rho_f/\mu_f)^{0.52} \quad [12]$$

It is seen that the scatter of the data is less than ±10 per cent (indicated by dashed lines), except for a few points around a Reynolds number of 80,000. At this flow rate, instability arose, presumably due to fluctuations in degree of thermal contact. This phenomenon will be the subject of later investigation. The exponent of 0.52 in Equation [12] compares with 0.6 for nonmetallic fluids flowing normal to tube banks in the same Reynolds-number range.

The results of two runs made with copper tubes at location No. 5 are shown in Fig. 5, along with the line from Fig. 4 for comparison. It is seen that the copper tubes with their excellent wetting, gave coefficients considerably higher than those for the chrome-plated tubes.

The wetted tubes were more sensitive to flow rate, giving a line of 0.66 slope as compared with 0.52 for the nonwetted chrome-plated tubes. At a Reynolds number of 50,000, the copper tubes gave a coefficient about 35 per cent greater than that for the chrome-plated tubes. This difference in coefficients between wetted and nonwetted tubes is in general agreement with the results of many investigators who have worked with liquid-metals systems.

The equation of the line, as drawn through the data points, for the copper tubes is

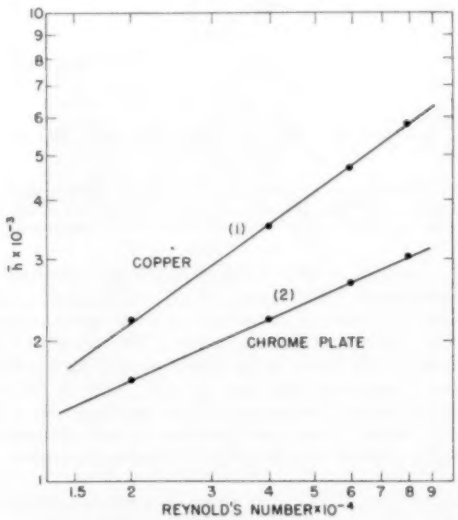


FIG. 6 COMPARISON OF AVERAGE HEAT-TRANSFER COEFFICIENTS FOR MIDDLE TUBE OF FRONT ROW UNDER WETTED AND NONWETTED CONDITIONS

$$\bar{h} = 3.45(D_o V_{max}\rho_f/\mu_f)^{0.66} \quad [13]$$

Front of Tube Bank. Only two runs were made on a tube in the front row, and this tube was in the middle of the row. One of these runs was made on a chrome-plated tube and the other on a copper tube. The results are shown in Fig. 6. In this case, the difference between wetted and nonwetted tubes was found to be greater than in the case of the interior tubes. At a Reynolds number of 50,000, for example, the front tube showed that wetting increased the coefficient about 65 per cent, whereas for the interior tube the increase was only about 35 per cent.

Under wetting conditions, the curve for the front tube fell slightly below that for the tubes in the interior of the bank; but

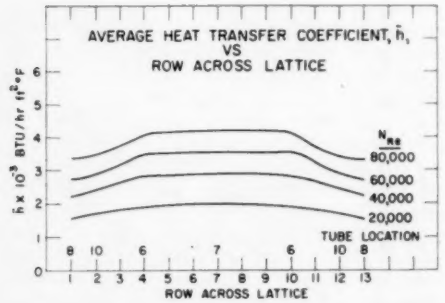


FIG. 7 VARIATION OF HEAT-TRANSFER COEFFICIENT $\bar{h}$ ACROSS TUBE BANK; I.E., PERPENDICULAR TO DIRECTION OF FLOW

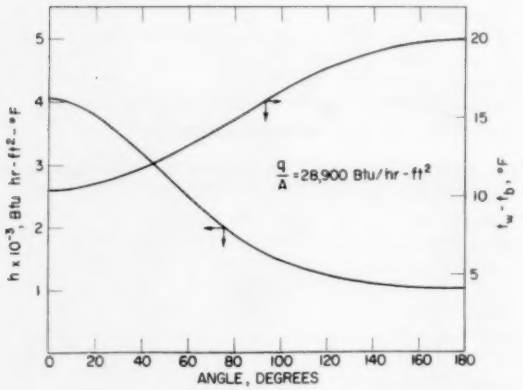


FIG. 8 VARIATION OF OUTSIDE TUBE-WALL TEMPERATURE, AS $(t_w - t_b)$, AND LOCAL HEAT-TRANSFER COEFFICIENT WITH ANGLE AROUND CHROME-PLATED TUBE FOR REYNOLDS NUMBER OF 2×10^4

under nonwetting conditions the front tube gave relatively much lower coefficients. For example, at a Reynolds number of 50,000 the front and interior chrome-plated tubes gave average coefficients of 2450 and 3600, respectively. This difference between front and interior tubes, in the case of nonwetting conditions, is in general agreement with experimental results on air (5) and water (6). No satisfactory explanation can be advanced for the fact that, under wetting conditions, the few data obtained on the front tube came so close to those for tubes in the interior of the bank.

Sides of Tube Bank. Two runs were made each at tube locations 8 and 10; the first being in the seventh transfer row from the front and the first row of full tubes from the left side, and the second being in the eighth transverse row and the second longitudinal row of full tubes from the right side of the bank. The two runs at each location checked very well. The results from these runs are summarized in Fig. 7 for four different Reynolds numbers. It is seen that the coefficient begins to fall off at about the third full-tube row from the side, the coefficient for the first full-tube row being about 80 per cent that for tubes in the central portion of the tube bank. This general result is what one would expect on the basis of the reduced flow near the sides of the lattice. To the authors' knowledge, these are the first results showing the transverse variation in the coefficient for tubes in a staggered tube bank where there are half tubes at the side walls.

Circumferential Variation of Coefficient. The variation of the local coefficient h around the circumference of a tube is shown, for typical cases, in Figs. 8 and 9. Values of h were calculated from

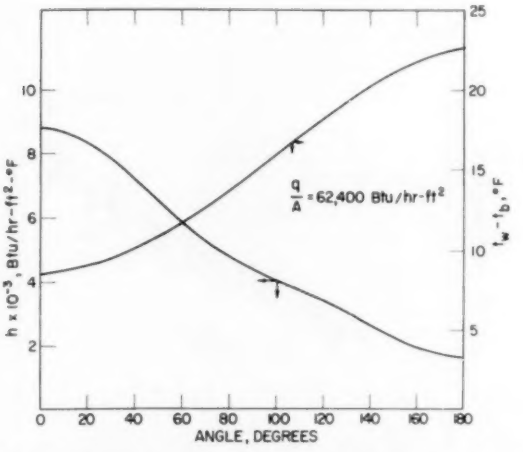


FIG. 9 VARIATION OF OUTSIDE TUBE-WALL TEMPERATURE, AS $(t_w - t_b)$, AND LOCAL HEAT-TRANSFER COEFFICIENT WITH ANGLE AROUND CHROME-PLATED TUBE FOR REYNOLDS NUMBER OF 8×10^4

CURVE	AUTHORS	FLUID	N_{Pr}	N_{Re}	P/D	SPACING
A	PRESENT	Hg	0.022	83,000	1.375	EQUILATERAL
B	De BORTOLI *et al*	AIR	0.70	74,000	1.30	EQUILATERAL
C	WINDING & CHENEY	AIR	0.74	60,000	SEE DIAGRAM BELOW	ISOSCELES

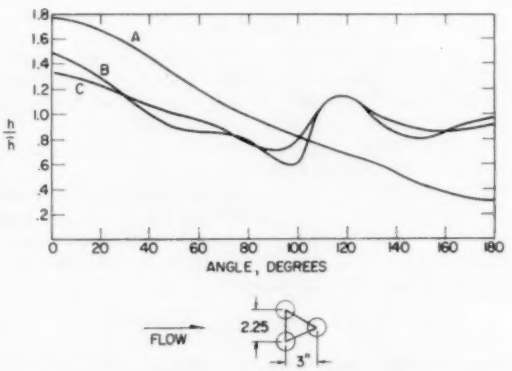


FIG. 10 RELATIVE VARIATION OF LOCAL HEAT-TRANSFER COEFFICIENT AROUND CIRCUMFERENCE OF A TUBE LOCATED IN INTERIOR OF STAGGERED TUBE BANK. TYPICAL RESULTS OF PRESENT STUDY ARE COMPARED WITH THOSE REPORTED ON AIR

the original data by means of Equations [10] and [11]. The figures also give the outside tube-wall temperature as a function of angle. Within the precision of the data, both sides of a tube facing downstream gave the same temperature profiles. In Fig. 8, where the results are for a Reynolds number of 2×10^4, there is a fourfold variation of the coefficient from the forward to the rear stagnation point. In Fig. 9, the variation is about fivefold for a Reynolds number of 8×10^4.

It is significant to point out that, if it had been assumed that there were no circumferential heat flow, the angular variation in the local heat-transfer coefficient would have turned out to be only about half that shown. These coefficient profiles are much smoother than those which have been obtained on air (7, 8) where weight losses from naphthalene tubes were measured and the mass-transfer/heat-transfer analogy (9) then used to estimate local heat-transfer coefficients.

Fig. 10 shows a comparison of present results with those reported by previous investigators on air under otherwise comparable conditions. Both the air profiles show relatively high maxima at about 120–130 deg, actually about $^3/_4$ of that at the forward stagnation point. There is another important difference between the metal and nonmetal profiles; at the same Reynolds number, there is roughly twice the variation in the coefficient around the tube for the mercury as there is for air. This is presumably due for the most part to the great difference in the values of the Prandtl numbers for the two fluids. Conceivably, it also could be partly due to the very much greater mass-flow rate of the mercury for the same Reynolds number, tending to reduce the heat-transfer resistance at the front of the tube and increasing it at the rear. Although the air profiles have been calculated from mass-transfer data, it is difficult to see how this fact alone can account for the great difference.

Since the profile curves shown in Figs. 8 and 9 were based on data obtained on chrome-plated tubes, it is likely that the relative smoothness of the curves may be due, in part at least, to the existence of nonwetting conditions.

The large variation in the local coefficient, and therefore the temperature, around the tube circumference in the case of liquid metals in crossflow could impose serious stress conditions in the tube wall at the zero stagnation point; which means that the question of circumferential variation of the coefficient is very important in heat-exchanger design where there is crossflow of a liquid metal.

Pressure Drop. The pressure-drop results are shown in Fig. 11 where they are compared with those of other investigators on air (10), water (6), and oil (11). The average deviation of the data points from the line in the figure was about 5 per cent. The friction factor f was evaluated by the following conventional equation for crossflow

$$\Delta p = (4fV^2_{max}\rho N)/(2g_0) \quad [14]$$

It is seen that the present curve falls appreciably below the indicated common curve. This result, in part at least, may be due to the nonwetting conditions which existed in the tube bank.

Conclusions

The results of the present investigation lead to the following conclusions regarding the crossflow of mercury in a staggered tube bank:

1 The local heat-transfer coefficient varies smoothly from a maximum at the forward stagnation point to a minimum at the rear stagnation point in the Reynolds-number range 15,000 to 80,000, the magnitude of the variation corresponding to a factor of 4–5.

2 The average heat-transfer coefficient for a tube in the interior of the bank varies as the 0.52 power of the velocity for nonwetted tubes and 0.66 for wetted tubes.

3 For tubes located in the interior of the bank, wetted tubes give heat-transfer coefficients considerably greater than those for unwetted tubes; e.g., at a Reynolds number of 5×10^4 they are 35 per cent greater. For a tube located in the front row, the difference was found to be much greater, the corresponding figure being 65 per cent.

4 The lower coefficients obtained in the front row of the tube bank compared to those in the interior, for the nonwetted tubes, are in general agreement with results obtained with ordinary fluids.

5 Tubes located at the side walls give coefficients about 20 per cent below those for tubes in the interior of the tube bank.

Acknowledgments

The invaluable assistance of Mr. William M. Hansen in developing and making the special heat-transfer elements is gratefully acknowledged. Moreover, discussions held with Mr. Corwin L. Rickard, his assistance with the calculation of the local coefficients, and his review of the manuscript have proved most valuable.

CURVE	AUTHORS	FLUID	P/D
A	BERGELIN *et al*	OIL	1.5
B	GRIMISON	AIR	1.5
C	DWYER *et al*	WATER	1.6
D	PRESENT	MERCURY	1.4

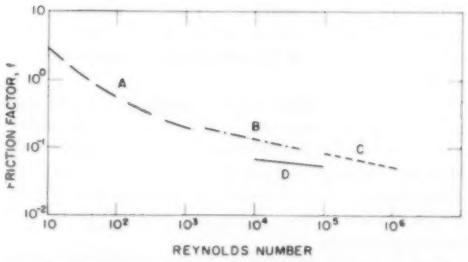


Fig. 11 Friction Factor Curves for Various Fluids Flowing Normal to Staggered Tube Banks With Equilateral Geometries

Bibliography

1 "Nuclear Power Plants: Design and Performance of Liquid Metal Heat Exchangers and Steam Generators," by R. O. Brooks and A. L. Rosenblatt, *Mechanical Engineering*, vol. 75, 1953, pp. 363–368.

2 "Performance of Small Liquid Metal Heat Exchangers," by R. A. Tidball, Preprints of Papers for Heat Transfer Symposium at Forty-Fourth Annual Meeting, the American Institute of Chemical Engineers, December 5, 1951, pp. 99–122.

3 "Heat Transfer and Pressure Drop in Heat Exchangers," by D. A. Donohue, *Industrial and Engineering Chemistry*, vol. 41, 1949, pp. 2499–2511.

4 "Liquid Metals Handbook," by R. N. Lyons, AEC and Bureau of Ships, Navy Department, second edition, June, 1952.

5 "Heat Transfer Symposium," by N. W. Snyder, the American Institute of Chemical Engineers Symposium Series, December, 1951.

6 "Heat-Transfer Rates for Crossflow of Water Through a Tube Bank at Reynolds Numbers up to a Million," by O. E. Dwyer, T. V. Sheehan, J. Weisman, R. T. Schomer, and F. L. Horn, to be published in *Industrial and Engineering Chemistry*.

7 "Mass and Heat Transfer in Tube Banks," by C. C. Winding and H. J. Cheney, *Industrial and Engineering Chemistry*, vol. 40, 1948, pp. 1087–1093.

8 "Average and Local Heat Transfer for Crossflow Through a Tube Bank," by R. A. DeBortoli, R. E. Grimble, and J. E. Zerbe, ASME Paper No. 55—SA-51.

9 "Mass Transfer Coefficients," by T. H. Chilton and A. P. Colburn, *Industrial and Engineering Chemistry*, vol. 26, 1934, p. 1183.

10 "Correlation and Utilization of New Data on Flow Resistance and Heat Transfer for Cross Flow of Gases Over Tube Banks," by E. D. Grimison, Trans. ASME, vol. 59, paper Pro—59-8, 1937.

11 "Heat Transfer and Fluid Friction During Flow Across Banks of Tubes," by O. P. Bergelin, G. A. Brown, and S. C. Doberstein, Trans. ASME, vol. 74, 1952, pp. 953–959.

12 "Pipe Friction Factors for the Turbulent Flow of Lead-Bismuth Eutectic," by H. A. Johnson, J. P. Hartnett, and W. J. Clabaugh, Report to U. S. Atomic Energy Commission from University of California, Berkeley, Calif., February, 1954.

Discussion

R. J. Grosh[6] and R. D. Cess.[7] The authors have made a

[6] Assistant Professor of Mechanical Engineering, Purdue University, Lafayette, Ind.

[7] Westinghouse Research Laboratory, Pittsburgh, Pa.

valuable contribution, both from the practical and theoretical viewpoints, to the problem of heat transfer to a cross-flowing liquid metal.

It is not clear whether the authors calculated the over-all average coefficient, using Equation [6], by assuming constant heat flux or whether it was assumed variable as in the calculation for the local coefficient of heat transfer. If the condition of constant heat flux was assumed, the question would arise as to whether or not the calculated values of the over-all coefficient would be representative of the actual case in which circumferential conduction occurs.

MacDonald and Quittenton[8] have suggested that the apparent wetting effect which has been observed is due to mechanical gas entrainment in the liquid metal. A number of investigators have reported findings concerning the effect of wetted and nonwetted conditions on heat transfer; the findings show that there is considerable disagreement. Were any precautions taken to reduce gas entrainment?

Fig. 10 of the paper gives a very good indication of the effect of the turbulent wake behind a cylinder upon heat transfer to liquid metals. For curves B and C representing the flow of air, it can be seen that a marked increase in the local coefficients will occur for the wake region behind the tube, this increase being due to the effects of eddies and vortexes. On the other hand, there is no increase in this same region for curve A, which represents the flow of mercury. This would indicate that the turbulent effects within the wake upon heat transfer are small compared to thermal conduction within the fluid.

J. W. Mausteller.[9] This work on mercury in crossflow has shed much needed light on a gray area in heat transfer. The authors are to be congratulated on the caliber of the work and the careful correlation of results. Determination of local coefficients is always to be desired and variation of coefficient around the tube is interesting—an extension of effort on this factor would be most desirable.

It may be of some interest to point out results obtained with NaK-56 (56 wt per cent potassium) in crossflow.[10] This work was curtailed before some of the facets of the experimental technique could be checked, but the data comprise the only work on NaK in crossflow. Two exchangers with $^1/_2$-in-OD tubes were used, one with staggered $^5/_8$-in. triangular pitch and the other with $^5/_8$-in. square pitch. Twenty five tubes were arranged in five rows of five tubes each, which is a smaller array than normally used. Exposed tube length was $6^1/_8$ in., and all tubes were active with NaK-56 as the heat source inside the tubes.

Outside-film coefficients were determined from the over-all coefficients, inside-film coefficient based on Martinelli's work, and the tube-wall resistance. Reynolds numbers ranged from 3000 to 70,000 and temperatures from 420 to 900 F with corresponding Prandtl numbers of 0.0128 to 0.0068.

Film coefficients as a function of Reynolds numbers were

$$h = 7.11\ \mathrm{Re}^{0.68}\ \text{staggered pitch}$$

$$h = 2.88\ \mathrm{Re}^{0.74}\ \text{in-line square pitch}$$

There was a definite spread with Prandtl-number change in the staggered-pitch exchanger (h versus Re on a line drawn by sight through data) but none with the in-line square. Plotting h as a function of Peclet number closed the spread to give $h = 98.5\ \mathrm{Pe}^{0.79}$. The higher coefficient on Reynolds number for the in-line square arrangement would be expected because of the greater effect of increasing turbulence. The L/D is rather low and prediction of the inside-film coefficients may be prejudiced which would affect the slope of the data; however, with identical entry sections, and so on, the two exchangers should be directly comparable. The data fall in the region which might be expected.

J. M. Savino.[11] The work presented here provides needed information for the designer of liquid-metal heat exchangers of the crossflow staggered-tube type. Because of the scarcity of such data for liquid metals, designers have been forced to use the available information for nonmetallic fluids. The authors state that the data obtained with mercury would be applicable to other heavy metals because the thermal and physical properties are all quite similar. To use the average heat-transfer-coefficient results for liquid metals other than mercury, the data should have been presented in a more generalized form; e.g., the Nusselt number (N_{Nu}) versus the Peclet number ($N_{Pe} = N_{Re}\ N_{Pr}$). This would also facilitate comparisons with data for nonmetallic fluids.

In the comparisons of the circumferential film coefficients (Fig. 7 of the paper), the inclusion of the data of Dwyer et al.[12] for water through a bank of staggered tubes would have been very interesting. The tube bank was 10 rows wide and 20 rows deep, and the tubes were spaced in an equilateral-triangle arrangement with a pitch-to-diameter ratio of 1.58. The water temperature was 360 F. The ranges of pressure, Reynolds number, and Prandtl number were 380 to 410 psig, 28,000 to 1,000,000, and 0.93 to 0.97, respectively. Fig. 12, reproduced from the reference, is for a cylinder in the central section of the tube bank, and it is seen that the coefficients at the rear stagnation point were about 80 per cent of those at the front stagnation point. The results for tubes near the side walls were essentially the same as for those in the center of the lattice. For the front row of tubes, the coefficients at the rear stagnation point exceeded those at the front by 15 per cent. Admittedly the conditions for the water tests were not comparable to the results shown in Fig. 7. However, it is significant to note that the results of the water tests do not resemble those of air even though the Prandtl numbers are of the same order of magnitude. If the water data were replotted using $h/\bar{h}$ instead of h for the ordinate, the curves would all lie in the range $1.2 < h/\bar{h} < 0.8$, Fig. 13. The curves for $N_{Re} < 230{,}000$ would be essentially straight and very nearly horizontal, while the curves for $N_{Re} < 230{,}000$ would retain the characteristic increase and decrease in $h/\bar{h}$ for increasing angle. It is important to note the sharp contrast in the variation and in the general shape of curves for the three fluids, particularly for $N_{Re} < 100{,}000$. It appears from this comparison that the differences between the data for air, water, and mercury cannot be attributed simply to the variation in Prandtl numbers.

The data related to the effect of wetting on the heat transfer was viewed with interest because the wetting problem has not been resolved and there is a need for more studies. Purely from intuitive reasoning, it seems that the type of wetting is one of the important factors which may affect the heat transfer. For example, wetting by the addition of an agent to reduce the surface tension is decidedly different from wetting by amalgamation; as

[8] "A Critical Analysis of Metal 'Wetting' and Gas Entrainment in Heat Transfer to Molten Metals," by W. C. MacDonald and R. C. Quittenton, Preprint No. 8 for Heat Transfer Symposium, American Institute of Chemical Engineers, December, 1953.

[9] Mine Safety Appliances Company, Callery Plant, Callery, Pa.

[10] "Heat Transfer and Pressure Drop With NaK-56 Flowing Perpendicular to Vertical Tubes," by M. J. McGoff and J. W. Mausteller, Memo Report 87 (BuShips Contract NObs-65426) Mine Safe Appliances Company, July 29, 1955.

[11] National Advisory Committee for Aeronautics, Lewis Flight Propulsion Laboratory, Cleveland, Ohio.

[12] "Heat-Transfer Rates for Crossflow of Water Through a Tube Bank at Reynolds Numbers up to a Million," Part 2—"Circumferential Variation of Film Coefficient for Individual Tubes," by O. E. Dwyer, T. V. Sheehan, and J. Weisman, ASME Paper No. 54—F-20.

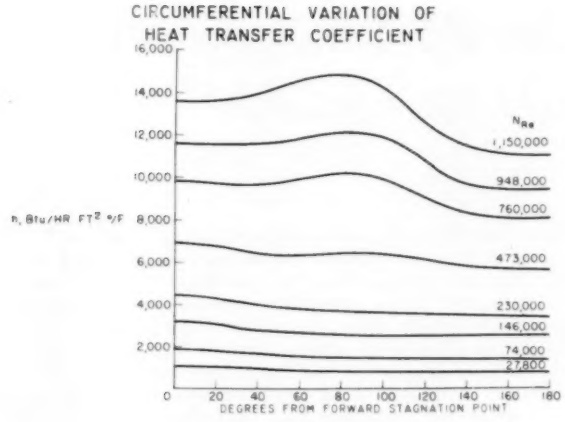


FIG. 12

CIRCUMFERENTIAL VARIATION OF HEAT TRANSFER COEFFICIENT

CURVE	AUTHORS	FLUID	N_{Pr}	N_{Re}	P/D	SPACING
A	PRESENT	Hg	0.022	83,000	1.375	EQUILATERAL
B	DEBORTOLI ET AL	AIR	0.70	74,000	1.30	EQUILATERAL
C	WINDING CHENEY	AIR	0.74	60,000		ISOSCELES
D	DWYER ET AL	H_2O	0.95	<250,000	1.58	EQUILATERAL

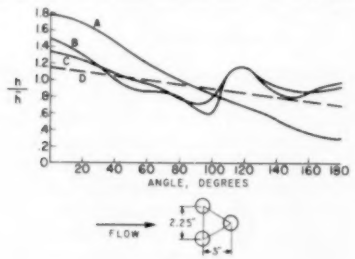


FIG. 13

a result, the thermal resistance at the interface between the wall and the liquid metal for each type of wetting may be different.

Other additional factors which would need to be considered are the cleanliness and roughness of the wall, the purity of the liquid metal, and the ambient temperature and pressure. Perhaps additional wetting studies could be incorporated into the heat-transfer program at Brookhaven National Laboratory. It would be of value to know how each of the two types of wetting would affect the local and average heat-transfer coefficients of the cylinder and the friction factor for the entire tube bank.

The mercury data in this paper show that the film coefficients increase with wetting by amalgamation. The authors stated that the bare copper tubes used for these tests were destroyed in two hours. It seems that the mercury attack upon the copper would soon roughen the cylinder surface, change the wall thickness around the cylinder, and change the thermocouple calibrations. Under these conditions, the validity of the results would be open to question. Whether or not the effects of mercury attack were present would depend on how soon the tests were run after installing the bare tube in the mercury and the length of time needed to make the runs. Some comment would be desirable explaining how the mercury-attack effects were avoided and how the reproducibility of the data was checked.

R. A. TIDBALL.[13] The authors are to be commended on the scope of the work reported and the unique experimental methods developed. It is particularly gratifying to see fundamental work presented in the field where so many unknowns exist.

It should be pointed out, however, that additional work on one shell-and-tube heat exchanger used to heat sodium in a segment-cut baffled shell was reported in the Sodium-NaK Supplement of the Liquid Metal Handbook, July, 1955. As the result of this work, Equation [2], presented in this paper was modified to

$$\mathrm{Nu} = 0.212(De)^{0.6}(\mathrm{Re})^{0.6}(\mathrm{Pr})^{0.6}$$

This equation varies only in that the Prandtl number is now raised to the 0.6 power instead of the $^1/_3$ power and is more consistent with liquid-metal heat-transfer data observed inside tubes.

It would be quite interesting if data were available from the present work to determine the effect of the Prandtl number on the heat-transfer coefficients in crossflow.

The writer also would be interested in hearing any comments on the variation of slopes for the wetted and nonwetted tubes.

[13] The Griscom-Russell Company, Massillon, Ohio.

AUTHORS' CLOSURE

Messrs. Grosh and Cess raise two questions. In reply to the first, a constant heat flux was assumed in the computation of average heat-transfer coefficients. It was found that if a variable heat flux were used, the values of $\bar{h}$ would have been higher by less than 4 per cent. To answer the second question, just normal precautions were taken to minimize gas entrainment. The effect of gas entrainment has been determined in a later study and will be presented in a second paper.[14] Under nonwetted conditions, the effect of slight gas entrainment on the coefficient was not great; while under wetted conditions, it was not detectable.

Mr. Mausteller presented some timely information concerning NaK crossflow heat transfer. However, we suspect that his reported coefficients are susceptible to significant error inasmuch as they were obtained indirectly by subtracting calculated tube-side and tube-wall resistances from measured over-all resistances.

We can appreciate Mr. Savino's desire to have the present results presented in the form of generalized equations, but it was not possible to do that here without making an assumption as to the effect of Prandtl number. The effect of Prandtl number has since been determined, and consequently the second paper to which reference is made[14] will include the generalized equations. Mr. Savino raised a matter which has not been satisfactorily explained, i.e., the shape of the coefficient versus angle profile curves obtained by Dwyer and co-workers at Brookhaven. It is felt that the relative flatness of these curves may be due, in part, to a certain amount of averaging of the tube-wall temperatures as measured; although this is not believed to be the chief explanation for the unexpected results.

Mr. Tidball raised the question of the effect of Prandtl number on the coefficient and suggested an exponent of 0.6 for this modulus in the modified Donohue equation. This is in line with recent results obtained at Brookhaven which showed that in an equation of the type

$$N_{\mathrm{Re}} = a + b\,(N_{\mathrm{Re}})^{c}\,(N_{\mathrm{Pr}})^{d}$$

the exponents c and d were the same and equal to 0.67.

Finally, the authors wish to express their appreciation to those who took part in this discussion, for their very pertinent questions and comments.

[14] "Heat-Transfer Rates to Crossflowing Mercury in a Staggered Tube Bank—II," by C. L. Rickard, O. E. Dwyer, and D. Dropkin, to be published.

Forces and Power Required to Turn Aluminum and Seven Alloys

By O. W. BOSTON[1] and W. W. GILBERT[2]

Turning tests on pure aluminum 1100-H14 and seven aluminum alloys were made to develop the formula for the tangential cutting force as a function of the material constant, the feed in inches per revolution, and the depth of cut in inches, when cutting dry, with a solid high-speed steel tool ground for turning aluminum. Equations for each metal have been developed and it is shown that the constants and exponents vary for each metal. Unit net power at the cutter has been computed for several sizes of cut for each metal and the values for a light cut and a medium-size cut have been plotted against each of the mechanical properties of the materials. These data show that knowing the Brinell hardness, the ultimate or yield strength of the metal or the shear strength, values of the unit net horsepower at the cutter can be computed with considerable accuracy.

Introduction

THIS paper presents the results of a number of tests dealing with cutting forces, and attempts to correlate them with the mechanical properties of the materials machined. The nominal chemical composition of the wrought-aluminum alloys received in 2-in-diam bars is given in Table 1. The mechanical properties of the pure aluminum and the seven alloys are given in Table 2 for each of the metals. In Table 2 the new temper designations for each alloy and temper are given, together with the old temper designations. The new designations were first published in the company's booklet.[3]

Cutting tools selected for these tangential-force tests were of an 18-4-1 type of high-speed steel in the form of solid bars, $^1/_2$ in. square as illustrated in Fig. 1. These were ground carefully to a tool designation of 20-deg back rake, 40-deg side rake, 10-deg end relief, 10-deg side relief, 10-deg end cutting-edge angle, 15-deg side cutting-edge angle, and a sharp nose. The bars cut to 24-in. lengths were clamped in the jaws of a chuck on the left end and supported on a live center on the right, in a 14-in. American Tool Works Company "Pacemaker" engine lathe. The lathe was driven by a 15-hp d-c motor powered from a Reliance Electric Company's motor-generator set to provide field and armature voltage control so that speeds from zero to 3000 rpm in infinite steps were available. This made it possible to machine the surface of any diameter at any desired cutting speed.

In the first series of tests, the tangential forces are measured with a tool dynamometer involving the S-4 strain gage and a Sanborn recorder. The force was determined for each of several speeds from 25 fpm up to 1000 fpm for each of the metals. The results shown graphically in Fig. 2 indicate that there is no appreciable reduction in tangential cutting force as the speed is increased. The slope of the curves is practically all the same, slightly lower to the right for the higher speeds, with a negative slope of 0.03 as indicated. For these cuts a constant feed of 0.0078 ipr and a depth of cut of 0.080 in. were used. All tests were run dry with the 20, 40, 10, 10, 10, 15, 0-in. tool of high-speed steel.

These tangential-force tests were continued for a constant speed of 100 fpm when the feed was varied for each of four depths of cut, and then the depth was varied for each of four values of feed. These results for the 1100-H14 aluminum are shown in Fig. 3. For the pure aluminum bars only, the tangential forces, as a function of feed, give a series of points lying on a curved line (dashed) for each of the four depths of cut. Straight solid lines, however, have been drawn to represent these data in order to have a single value of exponent for the feed. The slope of the lines for variable feed for each of the four depths of cut shown at the left in Fig. 3 is 0.60. This value represents the exponent of the variable f (feed) in the force equation. Similarly, at the right in Fig. 3, the slope of the lines for variable depth for each of the four feeds is 0.86. This indicates the equation

$$F_T = Cf^{0.6}\, d^{0.86}$$

By substituting the value of tangential force for given values of d and f, the constant is computed to be 5380, to give

$$F_T = 5380\, f^{0.6}\, d^{0.86}$$

as shown for the 1100-H14 aluminum in Table 3. Using this equation with the constant given, the tangential force for any other combination of feed and depth may be computed, or the values may be selected directly from the curves given in Fig. 3.

Similar tangential-force data, as a function of feed and depth when machining the 2024-T4 aluminum alloy at 100 fpm, are given in Fig. 4. In this case the slope of the force lines for the variable feed is 0.67, and that for the variable depth lines is 0.96. These values give rise to the formula

$$F_T = 17{,}900\, f^{0.67}\, d^{0.96}$$

shown for this alloy in Table 3. The constant has been computed to be 17,900 using the experimental value of F_T (61 lb) shown in Table 3.

Similar values of tangential cutting force for variable feeds and speeds were obtained for each of the other alloys. A summary of the tangential cutting forces, as a function of feed, for the constant depth of 0.080 in., for all of the alloys is given in Fig. 5. This shows that the 7075-T6 alloy requires about twice the cutting force as the 1100-H14 material. The lines for all the aluminum alloys are straight and nearly parallel (the pure aluminum 1100-H14, excepted). The slope of the variable feed lines for the alloys is represented by an average value of 0.7, which is the exponent of the variable, or $f^{0.7}$. This holds fairly well for all alloys except the 1100-H14 and 2011-T3.

[1] Professor of Mechanical Engineering and Production Engineering, Chairman of the Department of Production Engineering, University of Michigan, Ann Arbor, Mich. Fellow ASME.

[2] Consultant, Manufacturing Engineering Services Department, General Electric Company, Schenectady, N. Y.; formerly, Professor of Production Engineering, University of Michigan, Ann Arbor, Mich. Mem. ASME.

[3] "ALCOA Aluminum and Its Alloys," 1948.

Contributed by the Research Committee on Metal Processing and presented at the Semi-Annual Meeting, Cleveland, Ohio, June 17–21, 1956, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, February 8, 1956. Paper No. 56—SA-14.

TABLE 1 NOMINAL CHEMICAL COMPOSITIONS OF WROUGHT-ALUMINUM ALLOYS FOR MACHINING TESTS

(Per cent of alloying elements—aluminum and normal impurities constitute remainder)

Alloy	Copper	Silicon	Manganese	Magnesium	Zinc	Nickel	Chromium	Lead	Bismuth
1100	...	...	...	...	...	...	...	...	...
2011	5.5	...	...	...	...	...	...	0.5	0.5
2014	4.4	0.8	0.8	0.4	...	...	...	...	...
2017	4.0	...	0.5	0.5	...	...	...	...	...
2024	4.5	...	0.6	1.5	...	...	...	...	...
4032	0.9	12.5	...	1.0	...	0.9	...	...	...
6061	0.25	0.6	...	1.0	...	...	0.25	...	...
7075	1.6	...	0.2	2.5	5.6	...	0.3	...	...

TABLE 2 MECHANICAL PROPERTIES OF WROUGHT-ALUMINUM ALLOYS FOR MACHINING TESTS

New alloy, new temper	Old Alloy: New temper[a]	Old Alloy: Old temper	Ultimate strength, psi	Yield strength, psi[b]	Elongation, per cent in 2 in.	Reduction of area, per cent	Brinell hardness no.[c]	Shearing strength, psi[d]	Shearing strength / % elongation
1100-H14	2S-H14	2S-½H	17900	15300	35.5	68	32	10730	320
2011-T3	11S-T3	11S-T3	49400	38700	18.5	39	97	31870	1790
2014-T6	14S-T6	14S-T	71800	65000	13.0	25	139	46500	3580
2017-T4	17S-T4	17S-T	63400	42900	23.5	38	115	40670	1760
2024-T4	24S-T4	24S-T	68700	49700	19.0	26	122	41330	2170
4032-T6	32S-T6	32S-T	54500	48500	8.5	15	115	36430	4160
6061-T6	61S-T6	61S-T	43400	38800	19.5	51	94	29300	1500
7075-T6	75S-T6	75S-T	85100	76800	12.5	20	153	51230	4160

[a] Designations since January 1, 1948.
[b] Set = 0.2 per cent.
[c] 500-kg load on 10-mm ball. Average of tests at center, edge, and midway between.
[d] Determined from double-shear test.

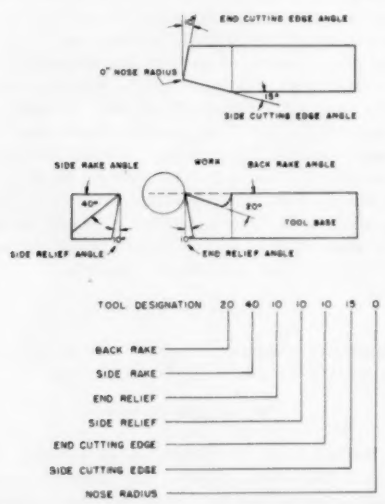


FIG. 1 NOMENCLATURE AND TOOL DESIGNATION FOR A TYPICAL SOLID HIGH-SPEED STEEL TOOL AS GROUND FOR TURNING ALUMINUM AND ITS ALLOYS

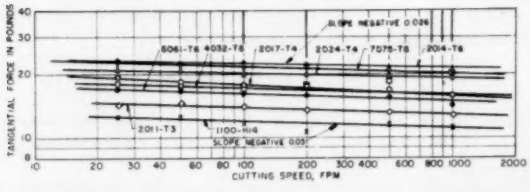


FIG. 2 INFLUENCE OF CUTTING SPEED ON TANGENTIAL FORCES WHEN TURNING ALUMINUM AND ITS ALLOYS, CUTTING DRY WITH A CONSTANT FEED OF 0.0078 IPR AND DEPTH OF CUT OF 0.030 IN.
(Tool of high-speed steel had a shape of 20, 40, 10, 10, 10, 15, 0-in. nose radius.)

In Fig. 6 is shown the relationship between the tangential cutting force and the depth of cut for a feed of 0.0078 ipr, when the cutting speed was 100 fpm. The slopes of these lines vary from a minimum of 0.86 for the 1100-H14, and 0.80 for the 2011-T3 to roughly 0.95 for the balance of the metals. These values represent the exponent of the variable depth and are summarized in the equations of Table 3, which shows also a general equation of $F_T = Cf^{0.7}d^1$, which is close for all metals except 1100-H14 and 2011-T3. The constants given in Table 3 should be used for each metal, however, for accurate values as was done in computing the values in Table 4.

The unit horsepower, u hp$_c$, that is, the net horsepower at the cutter per cubic inch of metal removed per minute is another means of representing the machinability of the aluminum and its

TABLE 3 EQUATIONS FOR TANGENTIAL CUTTING FORCES F_T AND VALUES OF "C" COMPUTED FOR EACH METAL USING TEST DATA INDICATED

(Tool shape 20, 40, 10, 10, 10, 15, 0-in. nose radius, and cutting speed 100 fpm)

Aluminum	F_T for cut 0.0078 f 0.080 d	Values of C and force equation $F_T = C f^b d^y$
1100-H14	33.1	$F_T = 5380\, f^{0.86}\, d^{0.86}$
2011-T3	38.5	$F_T = 8000\, f^{0.86}\, d^{0.86}$
2014-T6	61.5	$F_T = 20380\, f^{0.95}\, d^{0.97}$
2017-T4	54	$F_T = 18800\, f^{0.95}\, d^{0.99}$
2024-T4	61	$F_T = 17900\, f^{0.97}\, d^{0.96}$
4032-T6	54	$F_T = 17850\, f^{0.95}\, d^{0.97}$
6061-T6	49	$F_T = 13850\, f^{0.79}\, d^{0.99}$
7075-T6	66	$F_T = 23600\, f^{0.98}\, d^{1.02}$
All alloys[a]	..	$F_T = C\, f^{0.7}\, d$

[a] Approximate general equation for all alloys except 1100-H14 and 2011-T3 when using C for each metal.

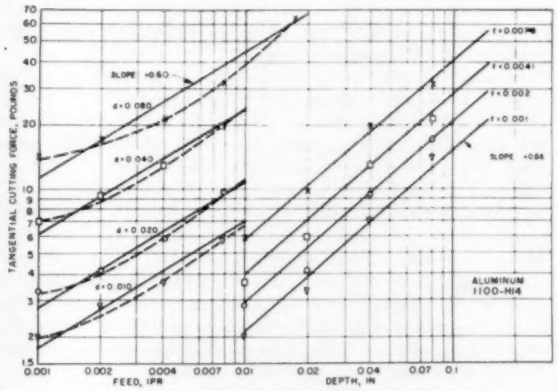


FIG. 3 TANGENTIAL CUTTING FORCES AS A FUNCTION OF FEED AND DEPTH OF CUT WHEN TURNING PURE ALUMINUM 1100-H14, DRY, AT A CONSTANT SPEED OF 100 FPM WITH STANDARD TOOL

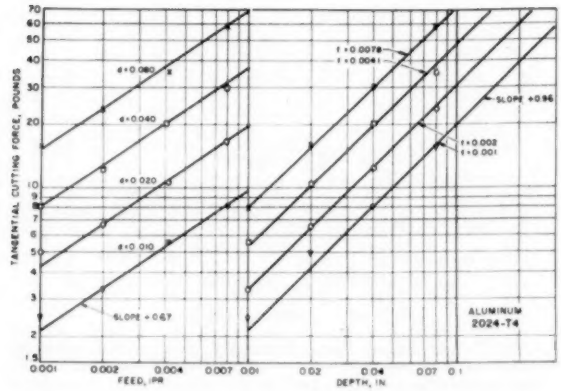


FIG. 4 TANGENTIAL CUTTING FORCES ARE SHOWN AS A FUNCTION OF FEED AND DEPTH OF CUT WHEN TURNING ALUMINUM ALLOY 2024-T4, DRY, AT A CONSTANT SPEED OF 100 FPM WITH STANDARD TOOL

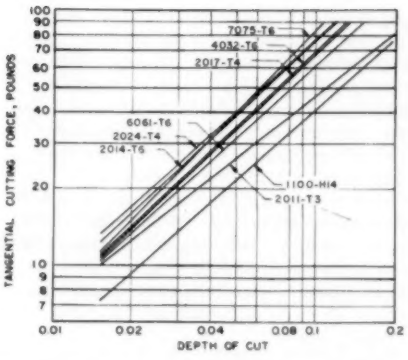


FIG. 6 TANGENTIAL CUTTING FORCE F_T VERSUS DEPTH OF CUT FOR FEED OF 0.0078 IPR WHEN TURNING ALUMINUM AND ITS ALLOYS AT A CONSTANT SPEED OF 100 FPM, CUTTING DRY

(Slopes of lines correspond to exponent of d in Table 3. Tool shape of 20, 40, 10, 10, 10, 15, 0-in. nose radius was used.)

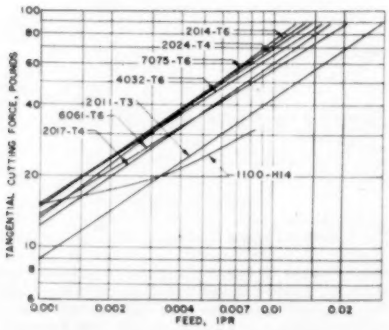


FIG. 5 RELATION BETWEEN TANGENTIAL CUTTING FORCE F_T AND FEED AT A CONSTANT DEPTH OF 0.080 IN. WHEN TURNING ALUMINUM AND ITS ALLOYS AT A SPEED OF 100 FPM, CUTTING DRY

(Slopes of lines correspond to exponent of f in Table 3. Tool shape of 20, 40, 10, 10, 10, 15, 0-in. nose radius was used.)

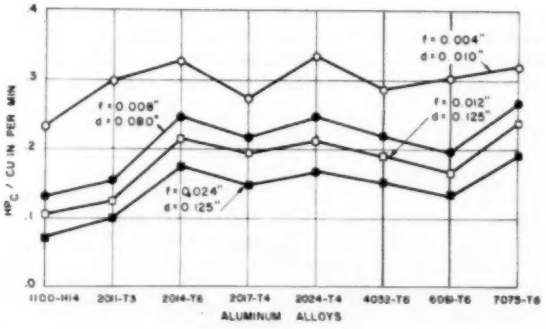


FIG. 7 VARIATION IN UNIT-HORSEPOWER REQUIREMENTS FOR ALUMINUM AND ITS ALLOYS WHEN TURNING DRY AT 100 FPM, FOR EACH OF FOUR CUTS SHOWN

(High-speed-steel tool with a shape of 20, 40, 10, 10, 10, 15, 0-in. nose radius was used.)

seven alloys. The net horsepower at the cutter is equal to the tangential cutting force F_T times the cutting speed V divided by 33,000. The unit horsepower at the cutter is equal to this value of hp_c divided by the cubic inches removed per minute. Therefore $u\ hp_c$ equals hp_c divided by $12\,f\,d\,V$, or

$$u\ hp_c = \frac{F_T V}{12 f d V \times 33{,}000} = \frac{F_T}{396{,}000\, f d}$$

For the 2014-T6 alloy, this becomes

$$\frac{127.6}{396{,}000\, f d}$$

For a feed of 0.012 in. and a depth of 0.125 in.

$$u\ hp_c = \frac{127.6}{396{,}000 \times 0.012 \times 0.125} = 0.215$$

as shown for this cut in Table 4 along with similar values for all metals for each of four sizes of cut, from a feed of 0.004 in. up to 0.024 in. These data show that the unit hp_c is lower for the values of heavier feeds. For example, for the 7075-T6 alloy, the unit hp_c is reduced from 0.318 for the cut of 0.004 × 0.010 to 0.189 for the cut of 0.024 × 0.125 in. This reduction is due principally to the increase in feed.

Values of $u\ hp_c$ are shown for each of the metals when taking a cut, dry, at 100 fpm for each of four sizes of cut in Fig. 7. For

TABLE 4 VALUES OF $u\ hp_c$ FOR EIGHT ALUMINUM METALS FOR EACH OF SEVERAL SIZES OF CUT AS COMPUTED FROM EQUATIONS OF TABLE 3

Metals	Exponents of f	Exponents of d	Values of F_T and $u\ hp_c$ for each cut: f = 0.004, d = 0.010		f = 0.008, d = 0.080		f = 0.012, d = 0.125		f = 0.024, d = 0.125	
			F_T	$u\ hp_c$	F_T	$u\ hp_c$	F_T	$u\ hp_c$	F_T	$u\ hp_c$
1100-H14	0.60	0.86	3.68	0.232	33.7	0.133	63.4	0.108	83.8	0.071
2011-T3	0.68	0.80	4.68	0.295	39.3	0.155	75	0.126	119.5	0.101
2014-T6	0.69	0.97	5.15	0.325	62.3	0.245	127.6	0.215	206	0.173
2017-T4	0.69	0.99	4.35	0.274	54.9	0.216	114.6	0.193	178	0.150
2024-T4	0.67	0.96	5.31	0.335	62.4	0.246	126	0.212	200	0.168
4032-T6	0.69	0.97	4.52	0.285	55.3	0.218	112.4	0.189	180	0.152
6061-T6	0.70	0.89	4.8	0.302	49.7	0.196	98.3	0.166	159	0.134
7075-T6	0.68	1.02	5.03	0.318	66.7	0.266	139.6	0.235	224	0.189

the heaviest cut, the highest value of u hp_c is for the 7075-T6 alloy. The next highest value is for the 2014-T6 alloy. The values for alloys to 6061 are nearly equal and still lower, but the lowest values are for 1100 and 2011. The greatest spread for the heaviest cut is from 0.071 for 1100-H14 to 0.189 for 7075-T6. The latter is 2.67 times the former. Further, the value of u hp_c for 7075-T6 for the lightest cut is 0.318 and it is 0.189 for the heaviest cut. The former is 1.68 times the latter.

The greatest over-all spread for all metals is 0.335 for the 2024-T4 alloy at the lightest cut to 0.71 for the 1100-H14 aluminum at the heaviest cut. This indicates the range or variation in net power at the cutter per cubic inch of metal removed per minute when cutting all aluminum metals at various sizes of cut in industry.

Influence of Various Mechanical Properties of Aluminum Metals on Unit Net Horsepower

To show the influence of the mechanical properties of the various metals studied in this paper on the unit net horsepower at the cutter, u hp_c, Figs. 8 to 13 have been prepared. In each case the value of the unit net horsepower is given as the ordinate and the mechanical property as abscissa. These figures are intended to show the relationships only in general terms. For example, in Fig. 8 the unit net horsepower is shown as a function of the Brinell hardness of the various metals for both a light cut and a medium cut. The values of power are taken from the highest curve and

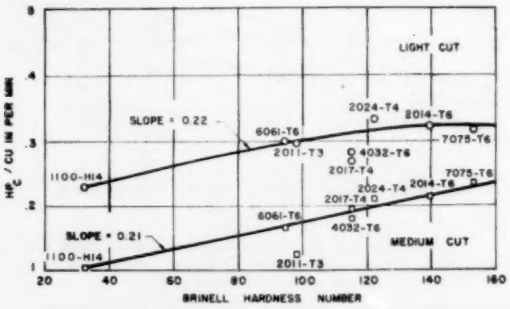


Fig. 8 Influence of Brinell Hardness on Horsepower Per Cubic Inch Per Minute When Making a Light Cut of $f = 0.004$ In. and $d = 0.010$ In., and a Medium Cut of $f = 0.012$ In. and $d = 0.125$ In. on Aluminum and Its Alloys at 100 Fpm, Cutting Dry

(Tool shape of 20, 40, 10, 10, 10, 15, 0-in. nose radius was used. $u\ hp_c$ = 0.00105 Bhn + 0.069 for the 0.012 × 0.125-in. cut.)

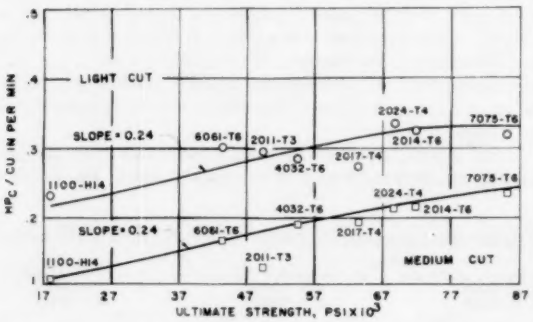


Fig. 9 Influence of Ultimate Strength on Unit Horsepower When Making a Light Cut of $f = 0.004$ In. and $d = 0.010$ In., and a Medium Cut of $f = 0.012$ In. and $d = 0.125$ In. on Aluminum and Its Alloys at 100 Fpm, Cutting Dry

(Tool shape of 20, 40, 10, 10, 10, 15, 0-in. nose radius was used. $u\ hp_c$ = 0.00000205 US + 0.071 for the 0.012 × 0.125-in. cut.)

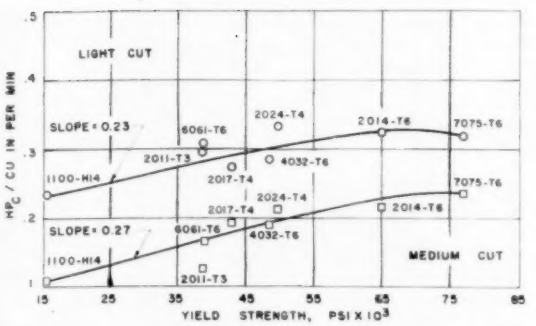


Fig. 10 Influence of Yield Strength on Unit Horsepower When Making a Light Cut of $f = 0.004$ In. and $d = 0.010$ In., and a Medium Cut of $f = 0.012$ In. and $d = 0.125$ In. on Aluminum and Its Alloys at 100 Fpm, Cutting Dry

(Tool shape of 20, 40, 10, 10, 10, 15, 0-in. nose radius was used. $u\ hp_c$ = 0.000002065 YS + 0.0764 for the 0.012 × 0.125-in. cut.)

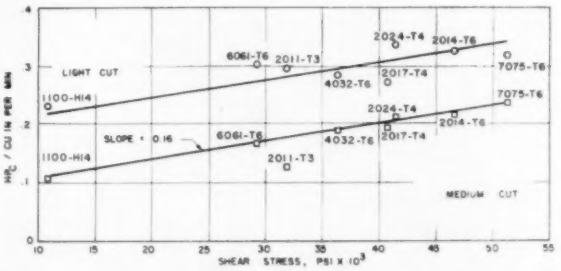


Fig. 11 Shear Stress in Pounds Per Square Inch Versus Unit Horsepower for Aluminum and Its Alloys for a Light Cut of $f = 0.004$ In. and $d = 0.010$ In. and a Medium Turning Cut of $f = 0.012$ In. and $d = 0.125$ In. at 100 Fpm, Cutting Dry

(Tool shape of 20, 40, 10, 10, 10, 15, 0-in. nose radius was used. $u\ hp_c$ = 0.00000314 SS + 0.0743 for the 0.012 × 0.125-in. cut.)

the third from the highest curve of Fig. 7, or from Table 4. For the heavy cut, which has a feed of 0.012 ipr and a depth of cut of 0.125 in., the relationship is almost a straight line. The unit power is increased from 0.108 to 0.235 (118 per cent) as the Brinell is increased from 32 to 153 (378 per cent). For the light cut, however, in which $f = 0.004$ in. and $d = 0.010$ in. a general relationship is indicated, although the values of power for the 4032-T6 and 2017-T4 are well below the indicated line while the value for 2017-T4 is slightly above. A line through the points for 2017-T4, 2014-T6, and 7075-T6 alone would show a negative slope.

The ultimate strength and its relation to the unit net horsepower at the cutter for each of the metals is shown for the light cut and medium-sized cut in Fig. 9. The lines are drawn merely to represent the relationship of the values to a normal expectancy. Practically all points for the medium-sized cut lie on, or close to, the line. The ultimate strength for 2011-T3 is considerably below the line and out of order. The point for 2017-T4 is also slightly below the line, although it is in relatively close agreement to the expectation. For the 118 per cent increase in unit power, there is an increase from 17,900 psi for 1100-H14 to 85,100 psi for 7075-T6, or 376 per cent. A greater deviation from the indicated line is shown for the value for the light cut, however. The point for 6061-T6 is slightly above the indicated line and the point for 2017-T4 is somewhat below the line. In fact, a line drawn through the points for 6061, 2011, 4032, and 2017 would be quite different

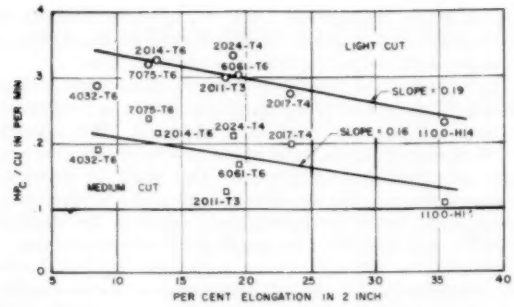


FIG. 12 HORSEPOWER PER CUBIC INCH PER MINUTE VERSUS PER CENT ELONGATION WHEN TURNING ALUMINUM AND ITS ALLOYS AT A LIGHT CUT OF $f = 0.004$ IN. AND $d = 0.010$ IN., AND A MEDIUM CUT OF $f = 0.012$ IN. AND $d = 0.125$ IN. AT 100 FPM, CUTTING DRY
(Tool shape of 20, 40, 10, 10, 10, 15, 0-in. nose radius was used.)

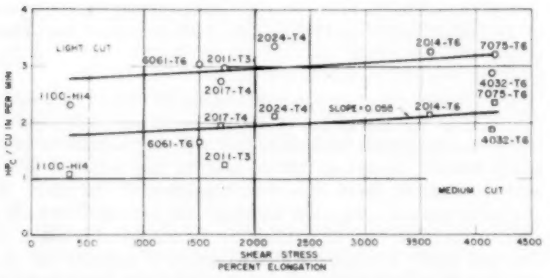


FIG. 13 INFLUENCE OF THE RATIO OF SHEAR STRESS TO PER CENT ELONGATION ON HORSEPOWER PER CUBIC INCH FOR TURNING ALUMINUM AND ITS ALLOYS WITH A LIGHT CUT OF $f = 0.004$ IN. AND $d = 0.010$ IN., AND A MEDIUM CUT OF $f = 0.012$ IN. AND $d = 0.125$ IN. AT 100 FPM, CUTTING DRY
(Tool shape of 20, 40, 10, 10, 10, 15, 0-in. nose radius was used.)

from that indicated for the light cut and have a negative slope indicating a reverse ratio.

Corresponding values of yield strength for the light and medium cuts, as a function of the unit power at the cutter, are shown in Fig. 10. The point for the free-cutting alloy, 2011-T3, is considerably below the line for the medium-sized cut. Also, the points for 2017-T4 and 2024-T4 are slightly above the indicated line. Otherwise, there appears to be a fairly direct relationship between the yield strength and the unit horsepower for the medium cut. The values for the light cut, as represented by the upper line in Fig. 10, show a greater dispersion from the indicated line. The two low points are for 2017-T4 and 4032-T6, while the points for 6061-T6 and 2024-T4 are well above the indicated line.

In Fig. 11 is shown the relationship of unit power to the shear stress as determined from a double-shear test. Except for the low value of power for 2011-T3, a single straight line seems to represent the straight-line relationship very well for all the metals for the medium cut. For the light cut several of the points are well off the indicated line indicating a less definite relationship between shear stress and unit power.

The per cent elongation is shown as a function of unit net power in Fig. 12. In this case, except for the value for 4032-T6, the points lie on an indicated straight line fairly satisfactorily for the light cut. A greater dispersion of the points from the indicated straight line for the medium cut is shown.

The shear stress divided by per cent elongation is the mechanical properties represented in Fig. 13 as a function of unit net power. The points for 1100-H14, 2011-T3, and 4032-T6 are well below the indicated line for the medium cut. The point for 7075-T6 is high. For the light cut, the points for 1100-H14, 2071-T4, and 4032-T6 are all well below the indicated line whereas 2024-T4 is well above the line. However, there does appear to be a trend for higher unit net power for higher values of shear stress over per cent elongation for both cuts.

CONCLUSIONS

When turning dry these several aluminum alloys with the high-speed-steel tool shape indicated as 20, 40, 10, 10, 10, 15, 0, the following general conclusions have been reached:

1 All metals give cutting-force values corresponding to exponential equations involving feed and depth, such as $F_T = Cf^x d^y$. However, each metal has its own peculiar exponents, x and y. In cutting most steels, for example, one set of exponents of feed and depth are obtained, and only the constant will vary for each steel. The aluminum alloys seem to be peculiarly individual in this respect.

2 In turning all eight aluminum metals at speeds from 25 to 1000 fpm, the cutting force remains practically constant for each metal. In other words, in this range of speeds there appears to be no marked variance in the cutting forces for the different metals.

3 The unit horsepower, that is, the horsepower at the cutter per cubic inch of metal removed per minute, varies almost directly with the Brinell hardness number of the metals for medium-sized cuts. The free-cutting alloy, 2011-T3, is well below the normal line, however. As the unit power is increased 118 per cent, the Brinell is increased 378 per cent. The equation for this line, so power may be computed from Brinell hardness, is $u\ hp_c = 0.00105\ Bhn + 0.069$.

Example: To determine the unit net horsepower at the cutter, $u\ hp_c$, if the Brinell hardness (Bhn) is known to be 94 (for the 6061-T6, Table 2) $u\ hp_c = 0.00105\ Bhn + 0.069$. (This is the equation of a straight line, of the form $y = mx + b$.) Then $u\ hp_c = 0.00105 \times 94 + 0.069 = 0.09975 + 0.069 = 0.16875$ which corresponds to 0.166 for this medium cut, Table 4. For a light cut there is a greater fluctuation of points; the direct relationship holds for only five of the eight metals.

4 The ultimate strength of all metals, except 2011-T3, gives almost a straight-line relationship with the unit power for the medium-sized cut. For the increase of 118 per cent in unit power, there is an increase of 378 per cent in ultimate strength. The equation for this line is $u\ hp_c = 0.00000205\ US + 0.071$, so unit power can be computed from ultimate strength (US). The power for the 2011-T3 is low for its strength. For the light cuts the relationship is more erratic for the several metals.

5 The yield strength (YS) increases almost directly as the unit net power, except for the 2011-T3 alloy which has relatively low power and 2017-T4 and 2024-T4 which have relatively high values, for the medium-sized cuts. For the 118 per cent increase in power, the yield strength is increased 400 per cent. Values of $u\ hp_c = 0.000002065\ YS + 0.0764$. Again, the relationship for the light cut between unit power and yield strength is more erratic.

6 The shear stress (SS), as determined from the double-shear test, gives a very good straight-line relationship between unit power and stress, the low values for 2011-T3 being one exception, for the medium cut. For the 118 per cent increase in power, the shearing strength is increased 378 per cent. For the medium cut 0.012-in. feed and 0.125-in. depth with the tool shown in Fig. 1, $u\ hp_c = 0.00000314\ SS + 0.0743$. For the light cut this relationship is less consistent.

7 The per cent elongation does not give a satisfactory linear relationship to the unit net power for the medium-sized cut for the various metals, but a better relationship is shown for the light cut except for the value of 4032-T6, which is low.

8 The shear stress divided by the per cent elongation does not give an over-all satisfactory linear relationship with the unit net power at the cutter for either the light or medium cuts.

Acknowledgment

The authors wish to thank the Aluminum Company of America through Mr. E. S. Howarth, Chief, Metal Working Division of the Aluminum Research Laboratories, for making the material with the chemical and mechanical properties available to us for this series of tests. The work was done in the Production Engineering laboratories of the University of Michigan by Walter Noffke, Research Engineer, now with the Douglas Aircraft Company, Glendale, Calif.

Mr. C. S. Cheng, a graduate student, assisted in the preparation of drawings and in checking the data. His service is gratefully acknowledged.

Discussion

E. S. Howarth.[4] This paper is another substantial contribution by the authors to the science of cutting metal. The results of their machining tests on aluminum and seven of its wrought alloys confirm the fact developed in a previous less comprehensive investigation[5] that the good machinability characteristics of aluminum alloys continue at high cutting speeds. In the work reported by Templin, no machinability limitation was found for the 2024-T4 aluminum alloy at surface cutting speeds as high as 20,000 fpm.

The wrought aluminum materials investigated by the authors represent a very wide cross section of aluminum alloys, ranging from commercially pure aluminum in the one-half hard temper (1100-H14) through one of the higher strength alloys (7075-T6). The approximately five-fold strength ratio represented by these two materials is one of the best indications of the differences in characteristics of the materials which were machined. The authors are to be complimented on the accuracy with which the tests were conducted and the resulting fine correlations which they obtained.

In most cases the data are consistent and follow patterns which were anticipated or can be explained. The greatest departures were observed in the cases of 1100-H14 and 2011-T3. The former of these, being commercially pure aluminum, did not receive its strength through heat treatment but rather by being cold worked to the one-half hard temper. Additional working which took place during the machining operation would, no doubt, further strain harden this material in the cutting zone and the amount of such strain hardening would probably vary to a greater extent for this material than for any of the others. The 2011-T3 alloy, containing lead and bismuth, is a free-cutting alloy which would normally be machined with tools having little or no rake. The relatively large rake angles selected for the constant tool employed by the authors was a good general selection but such angles are not desirable for this free-cutting alloy. Although the authors do not describe the characteristics of the chips which resulted in this investigation, I believe that under the large rake angle conditions this aluminum alloy, which is free cutting for zero rake angle, no doubt produced continuous chips.

The correlation between the unit net horsepower and the mechanical properties of the aluminum metals, as presented by the authors in Figs. 8 through 13, is very good. It appears that the shear strength of the material is a good criterion of the horsepower required for a specific cut. This might be anticipated because of the fact that the shear strengths were obtained from a double shear test which more closely simulates the machining operation than does either the tensile or Brinell hardness test.

Inasmuch as these machining tests were conducted under dry conditions, the frictional characteristics of the aluminum metals and their influence on machinability was probably greater than would be anticipated under conditions in which a lubricating-type cutting compound is employed. It would be interesting to extend the mechanical properties which the authors have presented in Table 2 to include coefficients of friction between these materials and high speed steel. It is believed that a consideration of these friction values would further explain the departure of 1100-H14 and 2011-T3 from the performance of the other alloys.

A. O. Schmidt.[6] This paper will be kept by many as a ready reference to the machining of aluminum alloys. The tests reported here are most interesting and valuable because of their completeness. Aluminum alloys used in the tests are quite representative of those commonly machined in the shop. It should be stressed again that the tangential cutting force hardly changes in value with an increase in cutting speed between 25 and 1000 fpm. The number of cubic inches of material that can be removed per horsepower may vary between 14 and 3, depending upon the type of cut and the type of alloy. Many times I have pointed out these variations to machine tool engineers and shopmen. Too often these factors are not given their proper consideration since there prevails a general notion that aluminum simply machines easily under any conditions. The correlation of physical properties to power required in machining, especially in regard to shear stress, is a definite part of the engineering groundwork in metal cutting. These data will help the machine designer and user to estimate properly the required machine capacity not only in turning but also in milling since face milling data correlate closely with those given here.

Authors' Closure

The authors appreciate the added information given by Mr. Howarth on the properties of the several aluminum metals, particularly the pure metal and the free-cutting type. Dr. Schmidt's comments, based on his experience, confirm the results given in the paper and add to its value. It is appreciated that tests of this nature on the machining of aluminum and its alloys could be extended to cover even higher cutting speeds, larger sizes of cut, different tool angles, and a variety of cutting fluids. We feel, particularly, that information on tool life would be of great value, and, of course, such tests should be extended to cover a variety of metal-cutting processes to complete the picture. These are suggestions for future work.

[4] Chief, Mechanical Engineering Division, Aluminum Research Laboratories, Aluminum Company of America, New Kensington, Pa. Mem. ASME.

[5] "Development of a High-Speed Lathe for Machining Aluminum," by R. L. Templin Trans. ASME, vol. 70, 1948, p. 837.

[6] Research Engineer, Charge of Metal Cutting, Kearney & Trecker Corporation, Milwaukee, Wis. Mem. ASME.

Comparative Machinability of B1113—C1213—C1120 HR—C1120 CD and C1119 Steels

By H. L. BRYDEN,[1] STORRS, CONN.

A report on extensive machinability tests from which C1119 steel indicated an average tool life six times greater than that of the best B1113 and C1213 steels and nine times the tool life of the average of seven varieties of B1113 and C1213 steels investigated.

Introduction

A SEARCH of the published data on machinability indexes for standard steels revealed practically complete standardization and agreement among all published tables when reduced to a common base.

Further study revealed almost identical procedure pursued in the technique used in establishing these machinability tables. Of all the published information uncovered, the ASME publication on "Manual of Cutting of Metals" was by far the most detailed and complete.

Lathe-turning tests are used for establishing the machinability ratings and are expressed in terms of tool-life expectancy at the three combined elements of speed, feed, and depth of cut. It is a splendidly devised system furnishing complete recommendations for just about every combination of speed, feed, and depth of cut on all standard steels to achieve the desired tool life. Tables are listed for 18-4-1 high-speed steel and carbide lathe tools of definitely specified shapes.

Selecting a group of steels among which were B1113 cold drawn and C1120 hot rolled, specifications of which are given in Table 1, an attempt was begun to establish the correctness of the ASME recommendations at selected rates of speed, feed, and depth of cut using the recommended HSS (18-4-1) lathe-tool shape.

TABLE 1 SPECIFICATIONS FOR CHEMISTRY—AISI—SAE CARBON STEELS

	Carbon	Manganese	Phosphorus	Sulphur
B1113.....	0.13 max	0.70/1.00	0.07/0.12	0.24/0.33
C1213.....	0.13 max	0.70/1.00	0.07/0.12	0.24/0.33
C1120.....	0.18/0.23	0.70/1.00	0.040 max	0.08/0.13
C1119.....	0.14/0.20	1.00/1.30	0.040 max	0.24/0.33

Lathe-tool shape No. 4, details of which are shown in Fig. 1, was selected for duplication and a depth of cut of $^1/_8$ in. and a feed of $^1/_{64}$ in. per revolution (ipr) were chosen as a starting point. Four-inch-diameter round bars 18 in. long were selected to insure rigidity in the workpiece. At the feed and depth of cut selected, using the No. 4 lathe-tool shape, the ASME tables were used to calculate the surface speed to give a 15-min tool life while cutting dry.

Reference to Tables 2 and 3, listing the results of these tests, will show the wide divergence experienced in the specific cases of

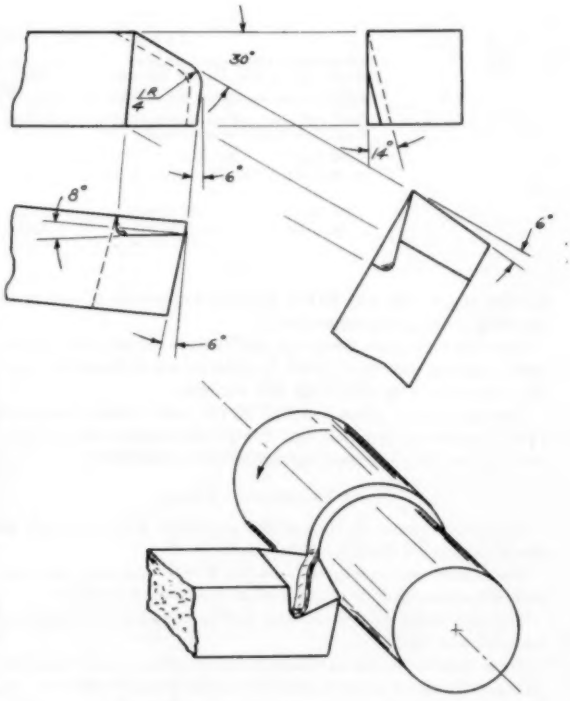


FIG. 1 ASME LATHE TOOL SHAPE NO. 4, 18-4-1 HSS

C1120 HR and B1113 CD from those predicted in the "ASME Manual."

Essentially what these tests showed was that B1113 CD, which is rated in the ASME classifications 30 per cent higher as a free-machining steel than C1120 HR, is actually only 80 to 85 per cent as free machining as C1120 HR. In other words B1113 is approximately 160 per cent overrated.

In most cases the end point in these lathe tests was the complete breakdown of the cutting tool which once begun is almost instantaneous. In the instances when complete failure was not achieved it was either lack of material at proper diameter to continue at specified surface speed or because nothing further would be established by continuing.

It is worthy of note here that all steels tested were first analyzed thoroughly to insure they were within the specification range and all lathe tools were prepared carefully, being formed by machine grinding before hardening and only lightly ground by machine after hardening just sufficient to remove the de-carb of hardening and insure precise tool angles being maintained. It also should be noted that several sources of supply were utilized

[1] Assistant Professor, University of Connecticut.

Contributed by the Research Committee on Metal Processing and presented at the Semi-Annual Meeting, Cleveland, Ohio, June 17–21, 1956, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, March 8, 1956. Paper No. 56—SA-26.

TABLE 2 SINGLE-POINT LATHE-TOOL TESTS

Actual feed, 0.0153 ipr; material, B1113 cold drawn; depth of cut, 1/8 in.; feed, 1/64 ipr cutting dry; tool shape, No. 4 (Fig. 1); tool material, 18-4-1 HSS

ASME recommendation for 15-min tool life = $\frac{1020\ HR}{178}$ (1.86)(1.19) = 394-fpm surface speed.

Each tool new and machine ground to identical geometry (hardness range 64 1/2–65 1/2 RC).
Tool life attained in minutes @; ( ) = tool No.; + = tool breakdown not reached.

183 fpm	224 fpm	227 fpm	243 fpm	247 fpm	267 fpm	270 fpm
22+(8)	5.10 (14)	10.03 (50)	4.22 (24)	25.14+(52)	3.33 (51)	0.72 (17)
...	18.01+(15)	17.03+(1)	...	...	3.84 (2)	...
...	16.29+(20)	...	...	...	17.41 (43)	...
...	10.97 (39)	...	...	...	4.94 (73)	...
289 fpm	**296 fpm**	**331 fpm**	**352 fpm**	**361 fpm**	**396 fpm**	**590 fpm**
1.86 (36)	1.52 (61)	0.67 (71)	0.29 (23)	0.82 (70)	2.35 (3)	0.065 (69)
0.33 (49)	2.29 (72)	...	...	...	0.17 (4)	...
0.60 (22)	0.55 (80)	...	...	...	0.19 (5)	...
...	...	...	...	...	...	...

TABLE 3 SINGLE-POINT LATHE-TOOL TESTS

Actual feed, 0.0153 ipr; material, C1120 hot rolled; depth of cut, 1/8 in.; feed, 1/64 ipr cutting dry; tool shape; No. 4 (Fig. 1), tool material, 18-4-1 HSS

ASME recommendation for 15-min tool life = $\frac{1020\ HR}{178}$ (1.42)(1.19) = 301-fpm surface speed.

Each tool new and machine ground to identical geometry (hardness range 64 1/2–65 1/2 RC).
Tool life attained in minutes @; ( ) = tool No.; + = tool breakdown not reached.

240 fpm	245 fpm	252 fpm	266 fpm	272 fpm	274 fpm	285 fpm
26.68+(10)	20.68+(46)	0.59 (13)	19.41+(21)	15.90+(30)	17.91+(60)	19.21+(18)
.....	.....	.....	.....	13.06+(40)	.....	.....
293 fpm	**297 fpm**	**306 fpm**	**308 fpm**	**311 fpm**	**333 fpm**	
21.56+(59)	13.14+(34)	1.71(42)	1.62(41)	1.48(45)	1.02 (56)	
.....	.....	....	....	3.70 (58)	....	

in selecting C1120 and B1119 material for testing as well as in choosing cutting-tool materials.

Once convinced the foregoing results were correct, the various steel suppliers were pressed for an explanation of the better tool-life experience with the C1120 HR material.

The explanation given was that B1113 in such large diameters (4 in.) was uncommon and that smaller diameters would produce results nearer the accepted standards for the two steels.

Small Diameters Checked

As a consequence of this, a test procedure was developed to check the smaller diameter (1 in.) steels.

These tests were conducted on a No. 2 wire feed-screw machine with all pertinent data as outlined at the head of Table 4.

The tool shape used is detailed in Fig. 2 and HSS 18-4-1 tool material was used.

Tools were to be run to complete breakdown or until tolerance on the diameter of pieces turned out varied beyond 0.0005 in. In every case the tool broke down before the tolerance was exceeded. Here again it should be noted that once tool breakdown begins complete destruction is almost instantaneous, and leaves no room for doubt as to when the end of a run is reached.

At the outset of these screw-machine tests it was intended merely to check C1120 versus B1113, having expected to return to accepted standards as predicted by suppliers once normal diameters were introduced. When discovered that C1120 was outperforming B1113 at the very minimum ratio of greater than 3 to 1 the question arose as to the grade of B1113 screw stocks being tested.

C1119 Introduced

At the same time an opinion was expressed that visibly the finish on the C1120 pieces was slightly inferior to those on B1113 although microinch tests indicated no difference.

Acknowledgment of this criticism led to the selection of C1119 steel, it being reasoned that if the B1113 appeared to produce a better finish, the additional sulphur content must be responsible. On the other hand the C1119 steel having practically identical chemistry to C1120, with the exception of added sulphur in the C1119, should continue the desirable machining qualities of C-1120.

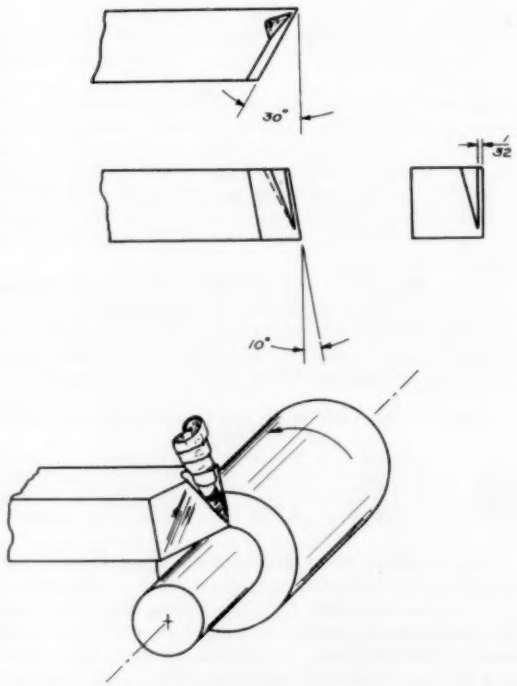


Fig. 2 Screw Machine Turning Tool 18-4-1 HSS

It was this introduction of C1119 steel showing even further improvement in desirable machining qualities as outlined in Table 4 and at the same time completely eliminating any further criticism as to finish, that the task of submitting all obtainable sources of B1113 and C1213 to the same test was undertaken.

In all, seven different B1113 and 1213 steels were obtained from a variety of suppliers and all subjected to identical treatment as outlined in Table 4 with results as recorded there. Table 4 is so

TABLE 4

DATA ON TEST RUNS MADE ON VARIETY OF SCREW STOCKS

[Single-point turning tests were conducted on a No. 2 wire feed-screw machine using standard grade cutting oil, to compare actual tool life on B1113 cold drawn; C1213 cold drawn; C1120 hot rolled; C1120 cold drawn; C1119 cold drawn. Using identical machining-ground HSS tool blades, each one new, and running to complete breakdown of tool or until tolerance varied beyond 0.0005 in. (In all cases the tool broke down before tolerance was exceeded.)] Length of cut, 3 1/2 in.; feed per revolution, 0.006 in.; depth of cut, 1/4 in.; stock size, 1 in.; rough turn from 1 to 0.975 in. (this was done on all materials tested for uniformity purposes because the hot-rolled stocks tested necessitated that scale first be removed); finish turn from 0.975 to 0.475 in.; spindle RPM, 1310-334-fpm surface speed (based on 0.975-in. stock diameter).]

Vendor A cold drawn B1113			Vendor B cold drawn C1213			Vendor C cold drawn C1213			Vendor D cold drawn B1113			Vendor E cold drawn B1113			Vendor F cold drawn C1213			Vendor G cold drawn C1213		
Test no.	No. pcs	Tl life, act min cut	Test no.	No. pcs	Tl life, act min cut	Test no.	No. pcs	Tl life, act min cut	Test no.	No. pcs	Tl life, act min cut	Test no.	No. pcs	Tl life, act min cut	Test no.	No. pcs	Tl life, act min cut	Test no.	No. pcs	Tl life, act min cut
1	97	43.6	1	90	40.0	1	41	18.4	1	55	24.5	1	116	51.6	1	53	23.6	1	58	
2	65	29.2	2	102	45.4	2	76	34.2	2	54	24	2	119	53.0	2	39	17.4	2	58	26.05
3	38	16.9	3	70	31.2	...	...	...	...	...	...	3	52	23.2	...	...	...	...	...	26.05
4	77	34.3	4	72	32.1	...	...	...	...	...	...	4	111	49.5	...	...	...	...	...	...
5	41	18.3	5	58	25.8	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
...	...	...	6	132	58.7	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
...	...	...	7	92	41.0	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
...	...	...	8	163	73.5	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Avg	63.6	28.5	Avg	97.4	43.5	Avg	58.5	26.3	Avg	54.5	24.25	Avg	99.5	44.3	Avg	46	20.5	Avg	58	26.05

GRAND AVERAGE ALL SCREW STOCKS 68.2 PIECES—30.5 MIN (1/2 HR)

Vendor	Hot Rolled C1120	Test no.	No. pcs	Tl life, min	Test no.	No. pcs	Tl life, min	Test no.	No. pcs	Tl life, min
H		1	321	145	2	350	156	3	545	243

GRAND AVERAGE C1120 HOT ROLLED 405 PIECES—181 MIN (3 HR)

Vendor	Cold Drawn C1120	Test no.	No. pcs	Tl life, min	Test no.	No. pcs	Tl life, min	Test no.	No. pcs	Tl life, min
J		1	325	145	...	...	...	...	...	...

GRAND AVERAGE C1120 COLD DRAWN 325 PIECES—145 MIN (2.4 HR)

Vendor	Cold Drawn C1119	Test no.	No. pcs	Tl life, min	Test no.	No. pcs	Tl life, min	Test no.	No. pcs	Tl life, min
K		1	625+	278	2	607	271	...	.	...

GRAND AVERAGE C1119 COLD DRAWN 616 PCS—274.5 MIN (4.58 HR)

NOTES: + = End point not reached.
C1119 = 6 times tool life of best screw stocks and 9 times tool life of average of all screw stocks.

arranged as to enable a rapid comparison of the average tool life experienced among all the varieties of steel stocks tested.

In consulting Table 4 it should be noted that all tests with the exception of the B1113 and 1213 steels were not run to their complete conclusions so that the claims of superiority are the very minimum and if all tests had been continued to their full conclusion the superiority would be even greater than here claimed.

Conclusion

A great many more tests were conducted than are recorded here both on lathe and screw machine using varied tool shapes in which the speed, feed, and depth of cut were varied, to ascertain that the superior machining qualities of the higher carbon steels (C1120—C1119) over the so-called free-machining lower carbon steels (B1113—C1213) maintained at the same approximate ratios.

The selected conditions under which the comparative tests were conducted were arrived at only after many preliminary trials and finally determined on the basis of obtaining the greatest possible amount of reliable information in relation to the time and effort expended.

From the findings here reported a criticism and an explanation are offered as to why such an error in accepted machinability ratings should have gone so long undetected.

The criticism lies (1) in the technique of conducting tests wherein tool pressures are measured as an indication of machinability and (2) that cutting tools are presumed to be exhausted on the basis of a predetermined amount of wear-land growth before complete destruction actually occurs. On the tests here reported measured tool pressures failed to indicate what tool life would be when carried to ultimate failure and in attempting to follow the wear-land growth technique an even less reliable relationship resulted. On the other hand in the technique here devised, running to complete breakdown of the tool and ignoring tool pressures and land growths, an end point is reached beyond any dispute and no hypothesizing or projecting is necessary; thus any prospect of error is eliminated from this source.

As an explanation of the results here reported the opportunity for the author to have observed so many tool failures due to the many tests conducted and the manner in which failure occurred, a pattern of failure eventually took shape of one type of steel versus the other. The failure patterns resolved themselves into two distinct classes wherein the steels of the higher carbon content (C1120—C1119) showed a continually steady increase in tool wear almost from the outset but stood up a great deal longer than possibly could have been anticipated when inspected microscopically at intervals during the life of the test. In the case of the steels of lower carbon content (B1113—C1213), microscopic inspection at similar intervals never revealed tool wear occurring as rapidly as in the case of higher carbon-content steels. What did occur time and time again in the case of the lesser carbon steels was unexpected failure when tool wear-land growth was much less than that of tools continuing runs on the higher carbon variety.

In short the tools wore out working the higher carbon steels and burned out unpredictably working the lesser carbon steels.

The explanation is thus advanced that as carbon content decreases the wear promoting carbide decreases and the free ferrite increases. This free ferrite in greater amounts is prone to weld to the tool with an attendant increase in frictional heat sufficiently great to burn the tool. Further the lower carbon steels tend to present a less homogeneous distribution of carbide and ferrite areas (banding), which creates a variable pressure on the tool as it moves alternately through carbon-rich and carbon-poor bands, and eventually, when an area of concentrated carbide particles is encountered, the heat created in surmounting this sudden demand is sufficient to destroy the tool instantaneously.

On the other hand in the case of the higher carbon steels the pattern of free ferrite to carbide areas is generally much more uniform and therefore presents a more even demand on the tool. It is thus suggested that an increase in carbon content at least up to 0.20 aids materially in improving machinability of a steel. Whether a further increase beyond 0.20 would be advantageous and how far beyond would be a very interesting study. A further advantage to an increase in carbon content beyond that presently contained in the B1113—C1213 grades over and above the improved machining qualities is the better hardening qualities inherent in the higher carbon steels.

Based on the author's recommendations, many production parts produced on screw machines have been turned out using C1119 steel at considerably elevated rates of production with enthusiastic reports of completely satisfactory results.

It is suggested respectfully that if new indexes were to be established for those set forth in the ASME tables, derived on the basis of the technique herein set forth, such tables would be the most informative, complete, and reliable of any that the author has yet encountered.

Discussion

S. L. Case.[2] The author's research on the comparative machinability of B1113—C1213—C1120 and C1119 steels is interesting. However, its main interest is not so much in the machinability ratings obtained on the steels tested as in the light it sheds on the hazards of accepting results of accelerated tests as a yardstick of any particular steel quality, be it machinability, corrosion, creep, or any other property reflecting performance under actual shop conditions, service-life span, and so on.

The accelerated machinability test employed by the author is widely used in laboratories. It furnishes a tool life measured in minutes instead of hours, as in actual performance in automatic machines. Results of accelerated tests, such as the single-point turning test, may be very informative if critically appraised, but should be accepted with reservations unless supported by large-scale tests on regular production in machine shops where many tons of metals are cut under each of many combinations of feed and speed. The author, himself, put his finger on the weak spot of his tests when he stated: "The failure patterns resolved themselves into two distinct classes wherein the steels of the higher carbon content (C1120—C1119) showed a continually steady increase in tool wear almost from the outset but stood up a great deal longer than possibly could have been anticipated when inspected microscopically at intervals during the life of the test."

The writer first observed this phenomenon more than 25 years ago during extensive machinability tests conducted on a four-spindle automatic lathe, under dry cutting conditions, so as to shorten the tool life. When this test furnished a relative machinability rating on a number of heats of a given grade, 3 tons of 1-in. hexagon bars from each heat were sent to a commercial machine shop and machined into bolts under carefully supervised conditions. So long as these tests were confined to bessemer grades B1112 and B1113, the machine-shop ratings were in line with the accelerated test ratings in the laboratory. However, on one occasion a low-sulphur open-hearth steel (SAE 1020) was tested in comparison with a B1112 heat. Surprisingly, under the accelerated test, the SAE 1020 heat showed a higher machinability rating than the B1112 heat, although the type of tool failure was different (gradual wear on the open-hearth heat, as compared with a sharp end point on the B1112 heat). When 3 tons of

[2] Technical Advisor, Battelle Memorial Institute, Columbus, Ohio.

cold-drawn bars from each of these heats were machined under normal conditions in the machine shop, the machinability ratings were found to be the reverse of those obtained under accelerated tests. Similar tests were made on a number of other heats (B1112—C1120—C1020) with the same end results. Eventually, this led to the adoption of different end point for the open-hearth steels—an end point based on a fixed value of tool wear, as compared with a sharp tool failure on the bessemer steels.

Incidentally, phosphorus and to a certain extent nitrogen, rather than carbon, appeared to be the cause of the difference in behavior of the steels under the accelerated machinability tests. Rephosphorized and nitrogenized (to about 0.012 per cent N_2) open-hearth steel behaved similarly to the bessemer grades, in so far as the end point of tool failure was concerned.

S. B. Jones.[3] The writer is very familiar with the machining qualities of C1119 steel. Over possibly five years, observations were made on many applications of the grade. The writer's interest in the results generally obtained—which often exceeded B1113 and lead-content steels—resulted in retaining the author to do some of the work on which he has reported.

It may be recalled that three or four years ago the author, Mr. George Witteman, and the writer appeared before Dr. Boston's committee in New York and pointed to some of our results. We said we felt the machining figures in the ASME Handbook needed revision. For one, the writer still believes they do, at least in the present instance.

Mr. R. F. Harvey, until recently chief metallurgist of the Brown and Sharpe Manufacturing Company, has published very significant findings with regard to tool life, operational speeds obtained, and general excellence of C1119. You will find his article in a fairly recent issue of *American Machinist*.

It is my belief that C1119 would be more generally used on difficult machine jobs were it not more costly to produce than conventional screw stock. The writer has been informed that the mills generally discourage the use of the grade for that reason.

[3] Precision Machine Tool Repairs Company, East Providence, R. I.

Author's Closure

Mr. Case's comment regarding the unreliability of accepting accelerated tests as a basis for production rates is one in which the author fully concurs. This very reluctance on the author's part to accept the results of the accelerated lathe tests was the reason for extending the tests to the screw machines. The basis of this paper is the projection of the accelerated lathe test results to actual production rates on screw machines. A tool life experience of $4^1/_2$ hours actual cutting time in the case of the C1119 steels is not considered by the author as an accelerated test in the case of normal screw machine operation. Automatic screw machine production setups normally call for 6 to 8 turrets or tools and in the most extreme case no single tool would be cutting more than 25 per cent of the production time per piece and in most cases much less. In an extreme case of an estimated 25 per cent cutting time for a particular tool, this would allow more than a 16-hour run between tool grinds which in the case of steel jobs is no more than the normal experience on production runs.

The surface speed rate of more than 300′/min for turning tools in the case of C1119 steels is fully recommended in actual production runs on screw machines which is fully 50 per cent in excess of most experiences uncovered by the author in the cases of production runs using B1113 and C1213 steels.

A note of caution regarding the hot-rolled steels, such as C1120, should anyone interpret this as their being recommended for screw machine runs at accelerated rates; do not attempt any such thing, as the type of chip produced from the hot-rolled steels will "clog" up at accelerated rates with the very good chance of a complete "smash up" resulting. For a rapid check of the accuracy of this data, it is recommended that for a production run on automatic screw machines, presently employing B1113 or C1213 steel, simply substitute C1119 steel of identical stock size, increase the speed 50 per cent, and continue the production run under these conditions.

Effect of Cold Work on Elevated-Temperature Properties of Types 301, 305, and 310 Stainless Steels

By R. A. LULA,[1] A. J. LENA,[2] and H. M. JOHNSON,[3] BRACKENRIDGE, PA.

The elevated-temperature strength of AISI Types 301, 305, and 310 stainless steels can be improved by cold-working, provided extensive recrystallization during exposure does not take place. The rupture strength of cold-rolled materials decreases progressively with increasing recrystallization and eventually becomes poorer than that of annealed material when the amount of recrystallization exceeds 50 to 60 per cent. It is suggested that microstructural changes which occur as a result of recrystallization account for the detrimental effect that recrystallization exerts on high-temperature strength. These include a change in grain size and redistribution of chromium carbides in Types 301, 305, and 310, and the accelerated formation of sigma in Type 310.

Introduction

THE ability of austenitic stainless steels to be work-hardened to high-strength levels has been used to great advantage for many years to provide corrosion-resistant, high-strength materials for structural components which operate at atmospheric temperatures. The outstanding example of this type of material is cold-rolled Type 301 stainless steel, which is used extensively for structural applications in automotive truck trailers and railway cars. In contrast to the wide use of the cold-rolled, austenitic stainless steels at normal temperatures, there has been relatively little attempt to extend their usage to elevated temperatures. This condition undoubtedly has been due to the lack of information describing the effects of cold-working on the strength and structural stability of these materials when subjected to loading at elevated temperatures. This paper is concerned with the results of an investigation on the effects of cold-working on the high-temperature properties of the austenitic stainless steels. The steels chosen for this work were Types 301, 305, and 310, with the major emphasis being placed on Type 301.

Materials and Procedures

AISI Type 301 was chosen initially for this study with additional work being performed on AISI Types 305 and 310 to provide wider coverage of the possible structural changes which can occur in the austenitic stainless steels at elevated temperatures. Type 301 transforms partially to martensite during cold-rolling, whereas, the other two grades are sufficiently stable to remain completely austenitic. Carbide precipitation in the temperature range of 1000–1600 F occurs in all three of these steels, but only in Type 310 does extensive sigma formation take place.

The composition of the particular three heats which were studied in this investigation is shown in Table 1. These steels were received from the mill in the form of annealed strip and portions of each strip were subsequently cold-rolled various degrees in the laboratory. The room-temperature properties of the three steels in the annealed and various cold-rolled conditions are given in Table 2. Samples of the annealed and cold-rolled material were studied with respect to structural changes including carbide precipitation, sigma formation, and recrystallization. Stress-rupture tests (2-in. gage length) were used to evaluate the effect of cold work on high-temperature properties.

TABLE 1 COMPOSITION OF STEELS

Type	C	Si	Mn	Cr	Ni	N_i	Cu
301	0.11	0.30	0.76	17.83	7.17	0.046	0.24
305	0.072	0.31	1.04	17.79	11.58	0.052	
310	0.065	0.35	1.40	24.58	19.86	0.050	0.16

TABLE 2 ROOM-TEMPERATURE TENSILE PROPERTIES

Condition, per cent	0.2% offset, yield strength, psi	Tensile strength, psi	Per cent elong in 2 in.
	Type 301		
As annealed	33330	110775	67.0
CR 5	61800	120350	46.0
CR 7.5	69350	127900	45.0
CR 10	71300	128725	44.0
CR 15	90375	150800	32.5
CR 20	111540	165000	27.5
CR 30	124300	166975	25.0
CR 40	178775	196000	9.0
CR 50	204150	217000	3.5
	Type 310		
As annealed	39860	86700	46.0
CR 5	62500	91870	38.0
CR 10	86795	102000	23.0
CR 40	126000	141800	8.0
	Type 305		
As annealed	26445	77850	77.0
CR 15	76200	94400	42.0
CR 30	109130	122490	20.0
CR 50	140530	155430	4.0

Results

Elevated Temperature Properties: Type 301. Conventional stress-rupture curves for annealed and cold-rolled Type 301 at 1200 and 1350F are shown in Fig. 1. The effects of cold work on the 100 and 1000-hr strength of this steel at 1200 and 1350 F can be seen in Fig. 2. The following generalizations can be derived from these curves:

1 Improvement in elevated-temperature strength derived from cold-working goes through a maximum at some intermediate degree of cold work, the exact value of which depends upon the temperature and time of testing. Maximum strengthening effect is obtained at 1200 F with 30 per cent cold-rolling for both 100 and 1000-hr rupture life. At 1350 F, very little improve-

[1] Research Supervisor, Stainless Steel Section, Research Laboratory, Allegheny Ludlum Steel Corporation. Assoc. Mem. ASME.
[2] Associate Director of Research, Research Laboratory, Allegheny Ludlum Steel Corporation.
[3] Metallurgist, Research Laboratory, Allegheny Ludlum Steel Corporation.

Contributed by the Metal Processing Research Committee and presented at a joint session of the Metal Processing Research Committee and Metals Engineering Division at the Semi-Annual Meeting, Cleveland, Ohio, June 17–21, 1956, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, March 21, 1956. Paper No. 56—SA-44.

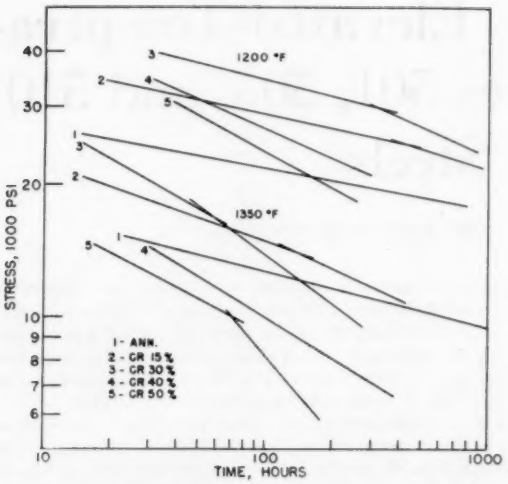


FIG. 1 STRESS-RUPTURE CURVES OF ANNEALED AND COLD-ROLLED TYPE 301 STAINLESS STEEL AT 1200 F AND 1350 F

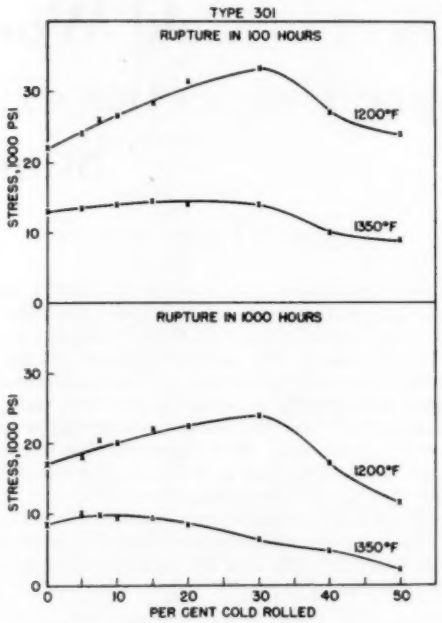


FIG. 2 EFFECT OF COLD-ROLLING ON STRESS-RUPTURE PROPERTIES OF TYPE 301 AT 1200 AND 1350 F

ment in 100 or 1000-hr rupture strength can be obtained by cold-working with large amounts of cold work (i.e., greater than 30 per cent for 100-hr rupture strength and 20 per cent for 1000-hr rupture strength) being detrimental to elevated-temperature strength.

2 For a constant-rupture life, the increase in strength derived from cold-working is greater in magnitude, the lower the testing temperature. For instance, a 50 per cent improvement in stress for rupture in 100 hr at 1200 F can be obtained by cold-working, whereas the maximum improvement at 1350 F for the same rupture life is about 10 per cent.

3 Excessive amounts of cold-working can result in rupture strengths less than that of annealed material with the minimum amount of cold work necessary to do this being dependent upon the temperature and time of test.

Cold-working has a detrimental effect on elongation at fracture in the stress-rupture test as shown in Fig. 3. An exception to this rule is found in samples which recrystallize, as shown by the 50 per cent cold-rolled material which had a low initial elongation at very short rupture times but the elongation increased rapidly thereafter coincident with a rapid rate of recrystallization.

Type 305. The stress-to-rupture curves for Type 305 at 1200 F are shown in Fig. 4. The 100 and 1000-hr rupture stress at 1200 F is shown in Fig. 5 as a function of the amount of cold work. These results are comparable to those of Type 301 at 1200 F in that 30 per cent cold-rolling produces maximum improvement for both 100 and 1000-hr rupture life. It is also interesting to note that the rupture strengths of the two alloys are essentially the same in spite of the difference in austenite stability. The stress to rupture in 100 hr of the 50 per cent cold-rolled Type

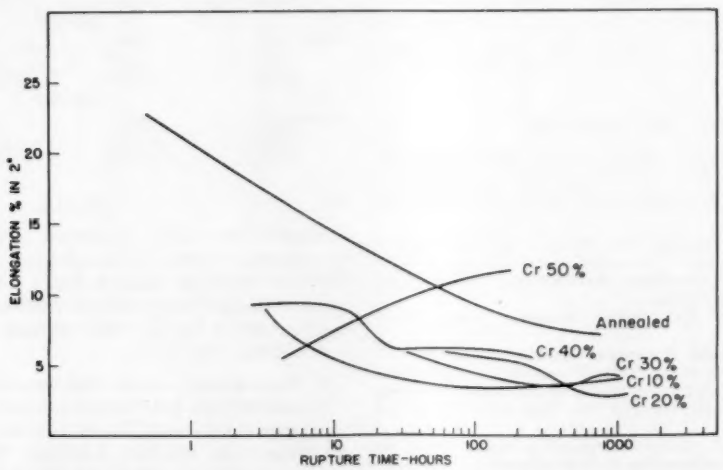


FIG. 3 EFFECT OF COLD-ROLLING ON ELONGATION AT RUPTURE IN TYPE 301 STAINLESS STEEL AT 1200 F

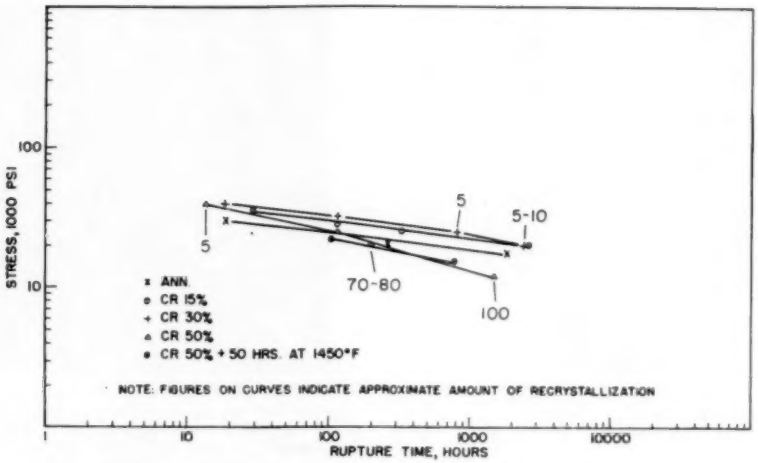


FIG. 4 STRESS-RUPTURE CURVES OF TYPE 305 AT 1200 F

305 is about the same as that for annealed material but is substantially less than that for annealed material for rupture in 1000 hr.

Type 310. The stress-to-rupture properties of Type 310 at 1350 F are shown in Figs. 6 and 7. Minor improvement at this temperature is obtained with 5 and 10 per cent cold-rolling with rupture lives of 100 and 1000 hr. The stress-rupture properties of material cold-rolled either 40 or 75 per cent is definitely inferior to that of annealed material for both 100 and 1000 hr. This effect is due to recrystallization which occurs very rapidly in heavily cold-rolled materials at high temperatures and has a detrimental effect on strength as will be described later. Information on the effect of intermediate degrees of cold work between 10 and 40 per cent was not obtained, but the data on Type 301 at 1350 F indicate that no further improvement in strength would be expected in this range.

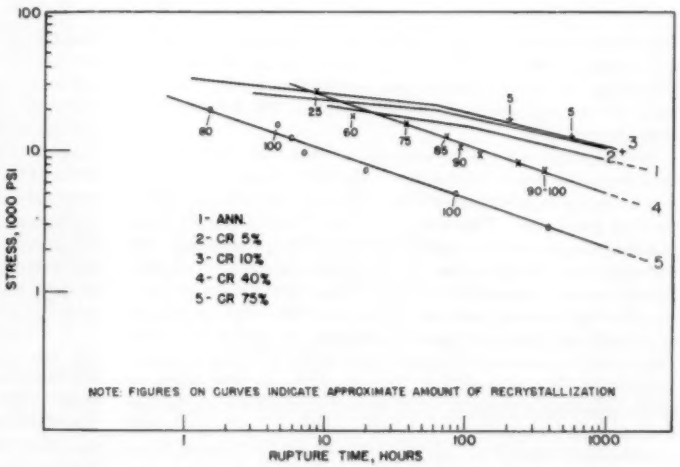


FIG. 6 STRESS-RUPTURE CURVES FOR TYPE 310 AT 1350 F

Structural Modifications at Elevated Temperatures. Metallographic examination of fractured test pieces has shown that recrystallization is the major structural change which affects the elevated-temperature strength of the cold-rolled austenitic stainless steels. When recrystallization begins, the rupture strength decreases and, in general, when recrystallization exceeds 50 to 60 per cent,

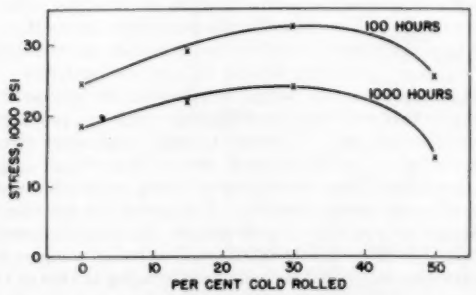


FIG. 5 EFFECT OF COLD-ROLLING ON STRESS TO RUPTURE IN 100 AND 1000 HR OF TYPE 305 STAINLESS STEEL

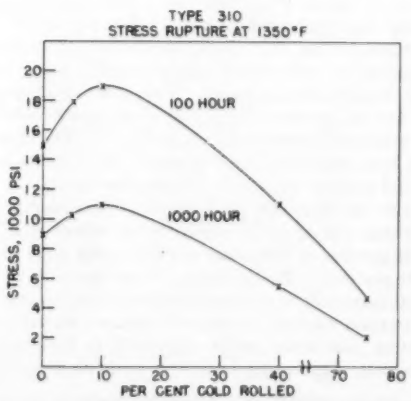


FIG. 7 EFFECT OF COLD-ROLLING ON STRESS TO RUPTURE IN 100 AND 1000 HR OF TYPE 310 AT 1350 F

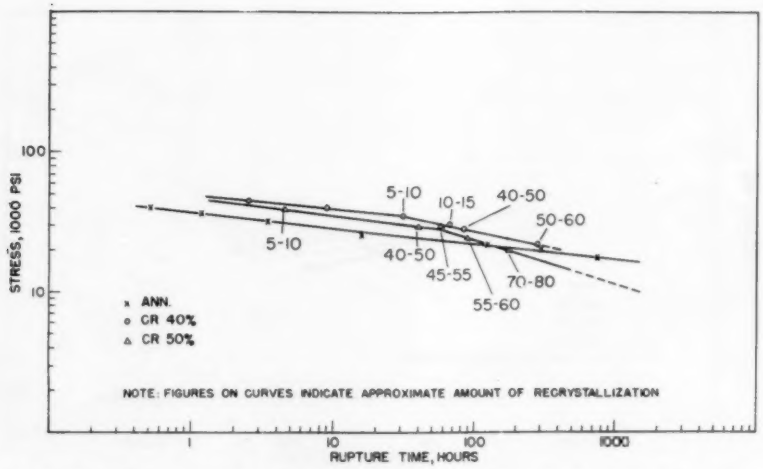


FIG. 8 STRESS-RUPTURE CURVES OF TYPE 301 AT 1200 F SHOWING EFFECT OF RECRYSTALLIZATION ON STRESS-RUPTURE PROPERTIES

the strength is less than that of annealed material. The effect of recrystallization can be seen in Figs. 4, 6, and 8, where the percentage of recrystallization for fractured test pieces is indicated on the stress-rupture curves for Types 305, 310, and 301, respectively.

Several structural modifications associated with recrystallization were observed in this study. The first of these, of course, was a change in grain size which in general resulted in a refinement of grain size. This was particularly true of heavily cold-worked samples which underwent rapid recrystallization and suffered rapid deterioration of elevated-temperature strength. Initial recrystallization in heavily cold-worked samples produced a high concentration of very fine grains at the boundaries of the cold-rolled steel, as can be seen in Fig. 9. In addition to the refinement in grain size, the photomicrograph of Fig. 9 indicates the partial disappearance of the chromium-carbide phase as recrystallization progresses. This resolution of chromium carbide during recrystallization is even more apparent in Fig. 10 which is a photomicrograph at ×2000 of 30 per cent cold-rolled Type 305 aged for 100 hr at 1350 F. The resolution of chromium carbide in this fashion was common to each of the three types of stainless steels.

Type 310 is unlike Types 301 and 305 in that it is capable of developing large quantities of sigma when exposed at elevated temperatures. Lena and Curry[4] previously showed that recrystallization of cold-worked samples during static aging accelerates the formation of sigma. This effect also was observed in fractured test pieces of Type 310 which had recrystallized in test, an example of which is shown in Fig. 11. The dark-etching phase in this photomicrograph is sigma and it exists in the recrystallized areas of a partially recrystallized sample.

Owing to the instability of Type 301, it transforms partially to martensite during cold-working to the extent that approximately 15 per cent of this phase is found in this alloy when cold-rolled 30 per cent. The presence of martensite is believed to contribute materially to room-temperature strengthening. In order to ascertain whether this phase is retained during stress-rupture testing and hence could contribute to high-temperature strengthening, dilatometric tests were made on 30 per cent cold-rolled Type 301 heated at a rate of 66 deg F per hr. Results of this test are shown in Fig. 12, where the inflection in the heating curve at 770 F establishes the temperature of transformation of martensite to austenite. The expansion coefficient between 1100 and 1400 F is the same as for annealed Type 301, indicating complete transformation. Hence, the presence of martensite at elevated temperatures cannot be a factor in the high-temperature properties of this grade. Further evidence for this is shown by the similarity in behavior between Type 301 and 305 in spite of the large difference in stability of the two steels. Type 305 when cold-rolled 30 per cent contains less than 0.5 per cent martensite, as determined by a magnetic test, whereas Type 301 in the same condition will contain approximately 15 per cent.

DISCUSSION OF RESULTS

The data presented in the previous sections of this paper have shown that it is possible to improve the elevated-temperature properties of the austenitic stainless steels by cold-working, providing the degree of cold work and conditions of exposure are not conducive to extensive recrystallization. Although the rupture strength begins to decrease with the initiation of recrystallization, improvement in strength is retained until the amount of recrystallization exceeds approximately 50 per cent.

An attempt was made in this study to determine why recrystallization exerts such a marked decrease in the stress-rupture properties of cold-rolled steels. One would expect that restoration of an annealed condition through recrystallization would produce strengths equivalent to steels tested in the annealed condition. However, the data presented in this paper have shown that complete recrystallization in cold-worked materials results in properties which are definitely inferior to annealed materials. This fact indicates that the actual recrystallization process during testing or that structural modifications which are produced by recrystallization are detrimental to high-temperature strength. With respect to recrystallization, *per se*, Becker[5] has suggested that accelerated creep should occur during recrystallization because of greater atomic mobility. This postulation has since been confirmed for a number of pure metals. In order to investigate this possible effect in cold-rolled stainless steels, samples of 305 and 310 were cold-rolled, recrystallized by aging at 1450 or 1475 F

[4] "The Effect of Cold Work and Recrystallization on the Formation of the Sigma Phase in Highly Stable Austenitic Stainless Steels," by A. J. Lena and W. E. Curry, Trans. ASM, vol. 47, 1955, p. 193.

[5] "Über Plastizität Verfestigung und Rekristallisation," by R. Becker, *Zeitschrift Technische Physik*, vol. 7, 1926, pp. 547–555.

FIG. 9 PHOTOMICROGRAPH SHOWING SMALL AMOUNT OF RECRYSTALLIZATION AT GRAIN BOUNDARIES IN STRESS-RUPTURE SPECIMEN OF 30 PER CENT COLD-ROLLED TYPE 305 WHICH FRACTURED IN 807 HR AT 1200 F WITH A STRESS OF 25,000 PSI. ELECTROLYTIC NITRIC-ACID ETCH—×500

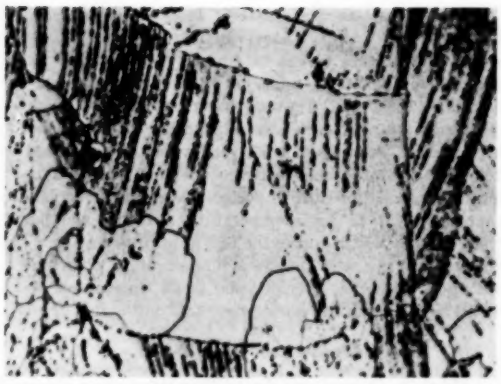


FIG. 10 PHOTOMICROGRAPHS OF COLD-ROLLED TYPE 305 AGED FOR 100 HR AT 1350 F SHOWING CARBIDE RESOLUTION. ETCHED ELECTROLYTICALLY IN SODIUM CYANIDE AND NITRIC ACID—×2000

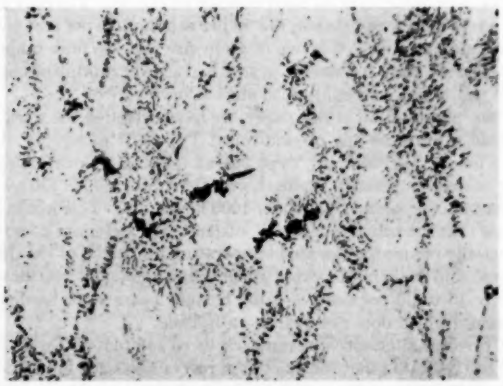


FIG. 11 PHOTOMICROGRAPH SHOWING SIGMA FORMATION IN RECRYSTALLIZED REGIONS OF 40 PER CENT COLD-ROLLED TYPE 310 STRESS-RUPTURE SPECIMEN WHICH HAD FRACTURED AT 1350 F IN 11.5 HR AT A STRESS OF 25,000 PSI. ELECTROLYTIC SODIUM-HYDROXIDE ETCH—×500

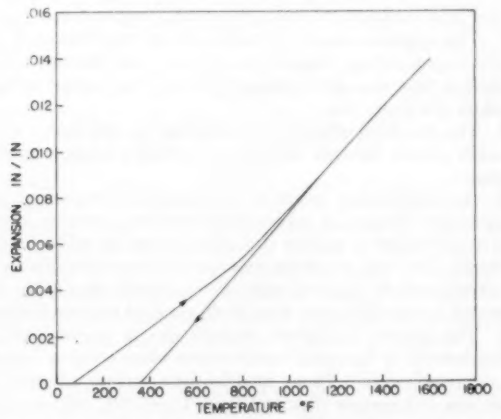


FIG. 12 DILATOMETRIC CURVE OF 30 PER CENT COLD-ROLLED TYPE 301. INFLECTION IN CURVE AT 770 F INDICATES TEMPERATURE OF TRANSFORMATION OF MARTENSITE TO AUSTENITE

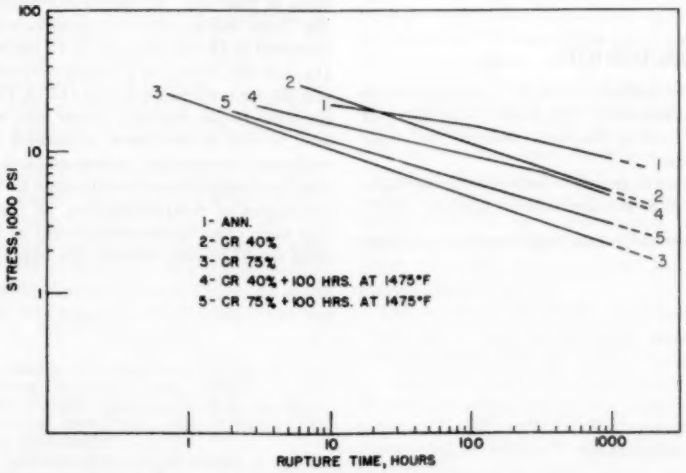


FIG. 13 STRESS-RUPTURE CURVES FOR TYPE 310 AT 1350 F SHOWING EFFECT OF AGING PRIOR TO TESTING

and subsequently tested at 1200 or 1350 F. The results are shown in the stress-rupture curves of Figs. 4 and 13. For both materials, the rupture curves of samples recrystallized prior to testing are well below the annealed curves for all rupture lives. Furthermore, the 1000-hr rupture strength of samples which were recrystallized prior to testing are substantially the same as those with equivalent amounts of cold work which recrystallized during test. As a result of these experiments it can be stated that the structural modifications which occur as a result of recrystallization are harmful to high-temperature strength whether the recrystallization occurs during the test or prior to testing. These microstructural changes include a refinement in grain size and modification of carbide precipitation in all three steels and accelerated sigma formation during recrystallization in Type 310.

Summary and Conclusions

1 Cold-working increases the rupture strength of Types 301, 305, and 310 stainless steel, providing the degree of cold-work and conditions of exposure are not conducive to extensive recrystallization.

2 The increase in strength derived from cold-working is greater in magnitude, the lower the testing temperature.

3 The optimum amount of cold-work for beneficiation is dependent upon testing temperature and time and decreases with increasing time at a given temperature or with increasing temperature at a given time.

4 The beneficial effect of cold-working on high-temperature strength persists through the recovery period prior to recrystallization.

5 Recrystallization results in a substantial decrease in high-temperature strength of cold-worked austenitic stainless steels. The improvement in rupture strength obtained by cold-rolling is gradually lost with increasing amounts of recrystallization, and the stress-rupture values of cold-rolled materials drop below the annealed values with more than 50–60 per cent recrystallization.

6 The decrease in rupture strength due to recrystallization is attributable to structural modifications which occur as a result of recrystallization. These modifications include changes in grain size and carbide distribution in Types 301, 305, and 310 as well as the accelerated formation of sigma in Types 310.

7 Martensite produced by cold-rolling in unstable austenitic stainless steels such as Type 301 transforms to austenite at testing temperatures and cannot contribute to high-temperature strengthening.

Discussion

T. L. Robertshaw.[6] The authors are to be congratulated on the fine work that is well presented in this paper. The discusser confesses there is nothing to add to this work; however, the paper has stimulated some questions:

(*a*) What was the difference in grain size between the specimens of 305 and 310 which were (i) as annealed and (ii) aged at 1475 F? Would such a difference account for the difference in stress rupture properties?

(*b*) Would the transformation of martensite to austenite occurring at 770 F in the cold worked 301 materially reduce its properties at 770 F?

(*c*) Did the amount of recrystallization vary along the gage length of the stress-rupture specimens?

(*d*) Did fractures "initiate" or propagate in a preferred manner so far as grain boundaries, precipitates, recrystallized grains, etc. are concerned?

[6] Research Assistant, Universal-Cyclops Steel Corporation, Bridgeville, Pa.

Authors' Closure

Examination of the Type 310, which had been aged at 1475 F, showed that the steel prior to aging had an extremely fine grain size which by calculation would amount to an ASTM grain size of 18–20. The annealed steel had a grain size of 5–6. This difference in grain size undoubtedly contributes to the decreased rupture strengths of aged samples. As a matter of fact, the authors consider the refinement in grain size possibly the most significant microstructural change which contributes to decreased rupture life when recrystallization occurs. The effect of grain size alone can be seen in the data for aged samples in Fig. 13. The only difference in structure between the 40 per cent and 75 per cent cold-rolled and aged steel was one of grain size and the finer grained 75 per cent cold-rolled steel (Curve 5) had a lower rupture strength than the 40 per cent cold-rolled steel (Curve 4).

Since presentation of this paper, we have extended the work to include rupture testing of cold-rolled Type 301 at 800 F. We have found that full hard Type 301 at this temperature retains its elevated temperature strength with the rupture life in 100 hours being 145,000 psi and 135,000 for 1000 hours' life. This would indicate that the transformation of austenite to martensite does not materially reduce the elevated temperature properties of the steel. This should probably be expected since the amount of martensite is only 15 to 30 per cent with the balance being work hardened austenite which does not change on heating.

A brief examination of several fractured samples of Types 301 and 305 indicated that the amount of recrystallization in partially recrystallized samples may have been slightly greater near the fracture. The difference was not great and was nonexistent at distances slightly removed from the fracture.

A number of the fractures in all grades of steel were examined since receipt of Mr. Robertshaw's discussion to determine the mode of fracture. All fractures examined failed intergranularly. For Type 305 at 1200 F this included annealed samples which fractured in 18 and 259 hours; 15 per cent cold-rolled steel in 29, 114, and 331 hours; 30 per cent cold-rolled in 18.5 and 114 hours; and 50 per cent cold-rolled in 112.75, 195.5, and 1508 hours. The annealed 15 per cent and 30 per cent samples had not recrystallized in test in the times mentioned whereas the 50 per cent cold-rolled sample had undergone partial recrystallization in short time and complete recrystallization in 1508 hours. Regardless of the degree of recrystallization, all fractures were intergranular. The same observations were found with 301 and 310. These results would further indicate the importance of grain size.

Avon No. 8—A Supercritical-Pressure Plant

By C. A. DAUBER,[1] CLEVELAND, OHIO

This paper describes the 250,000-kw supercritical-pressure addition to the Avon Plant of The Cleveland Electric Illuminating Company. Special features of the third supercritical-pressure unit designed for an American central station include the use of a monotube once-through boiler, combination motor and turbine-driven boiler feed pumps, extruded pipe for main steam leads, condenser by-pass demineralization system, and a high-temperature water-heating system for the plant addition.

Introduction and Economics

THE Cleveland Electric Illuminating Company has four generating plants and serves a 1700-square-mile area in northeast Ohio. This year the company is observing its 75th anniversary, its earliest predecessor having been founded in 1881 by the celebrated arc-lamp pioneer, Charles Francis Brush. Growth of both company and area is indicated in these words spoken during the official ground-breaking for the new unit by the company's President, Elmer L. Lindseth: "In 1881 the Illuminating Company's earliest ancestor, The Brush Electric Light and Power Company, had one steam generating unit which when run at full capacity would light a total of 88 lamps. When the new unit begins operation at Avon, the Illuminating Company will be able to turn out a total of two million kilowatts... This will be two and one-half times our generating capacity at the close of World War II."

This tremendous load growth has been a challenge to the company's engineers to design and construct the most efficient and economical generating plants. The recent breakthrough of the critical-pressure barrier for steam and water, while it provides a new tool for higher thermal efficiency, involves a host of intriguing problems in such fields as high-temperature metallurgy and boiler-feedwater conditioning. This paper is written to attract the interest of young engineers who may not be aware that electric power generation is an area of technology which has achieved, in the case of The Cleveland Electric Illuminating Company, a reduction in fuel consumption from 6 lb per kwhr to 0.67 lb per kwhr in the past 75 years. To these young engineers, as well as to those more experienced in central-station design and operation, it is hoped that the information will convey some of the challenge that faced engineers of our company when initial announcement was made in the summer of 1955 that Unit No. 8 for the addition to Avon Plant was to employ a supercritical-pressure steam cycle.

Since emphasis in this paper is on unusual features of design, many details duplicating conventional central-station practice will not be discussed. All aspects of the design had not been "frozen" at the time of preparation of the manuscript, so some portions may be subject to future revision. It was thought that a greater contribution to central-station practice could be made by sharing information at this time than by waiting until the station begins operation.

[1] Director of Civil and Mechanical Engineering Division, The Cleveland Electric Illuminating Company. Mem. ASME.

Contributed by the Power Division and presented at the Semi-Annual Meeting, Cleveland, Ohio, June 17–21, 1956, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, June 17, 1956. Paper No. 56—SA-69.

Extensive studies were carried out to establish the location of the 1958 addition, its optimum size, and the most economically desirable steam conditions.

Site location involved a choice between a recently purchased location in the Ohio River and several sites in various stages of development within the company service area. Strong consideration was given to the advice of the U. S. Defense Department to locate the unit outside of the Cleveland metropolitan area as a part of the policy of plant dispersal. By choosing to extend the Avon Plant, located 20 miles west of downtown Cleveland, it was possible to obtain a Certificate of Necessity for a five-year amortization on 65 per cent of the plant investment. Careful analysis of the cost factors involved indicated that the saving in coal-delivery charges for the river site were not as yet sufficient to overcome the added fixed and operating charges associated with a high-voltage transmission line from the Ohio River to the company service area. On the other hand, the Avon site did not involve comparable transmission-line costs, and it was found that the use of money made available by the deferment of taxes through rapid writeoff overcame all other economic factors and justified the selection of the existing station site on Lake Erie as the preferred location.

Selection of unit capacity size involved a study made in conjunction with the Operation Research Staff of Case Institute of Technology. An analysis of system reserve requirement, both with respect to maintenance and reliability, was carried out. This phase of the study resulted in the sizing of the unit at 250,000 kw and confirmed the in-service date of late 1958. The company established a forced-outage probability requirement of "no more than one deficiency in generating capacity in five years."

A number of heat cycles ranging in throttle pressure from 2400 psig to 5000 psig and in temperature from 1050 F to 1200 F were selected for consideration. Preliminary calculations narrowed the range to 3000–3500 psig and 1050–1100 F, with single reheat to 1050 F. Because of anticipated design and operating problems with steam at pressures just below the 3202-psia critical pressure, it was decided to move through this range to 3500 psig. Furthermore, the company's policy is to select steam conditions one step in advance of the most economical in order to contribute to the advancement of the art. In dealing with one turbine manufacturer, it was found that ferritic materials could be used to withstand 1100-F main steam temperature. For these reasons, throttle steam conditions of 3500 psig, 1100 F were selected with reheat to 1050 F.

Space limitations, together with electrical cost factors, dictated the selection of a tandem-compound unit. These considerations were influenced by the location of the new unit in an existing plant. The new extension to the Avon site will be as shown on the plot plan in Fig. 1.

The layout of the Avon extension is of the fully enclosed "ranch type" with the turbine-generator shaft oriented parallel with the front face of the steam-generating unit as shown in Fig. 3. This is the best arrangement, since the physical dimension of the tandem-compound turbine-generator along the center line of the shaft will be 108 ft. Walls for the station extension will be constructed of insulated stainless-steel panels with brick pilasters.

There will be a centralized control room for all boiler and turbine controls. This air-conditioned control room will be de-

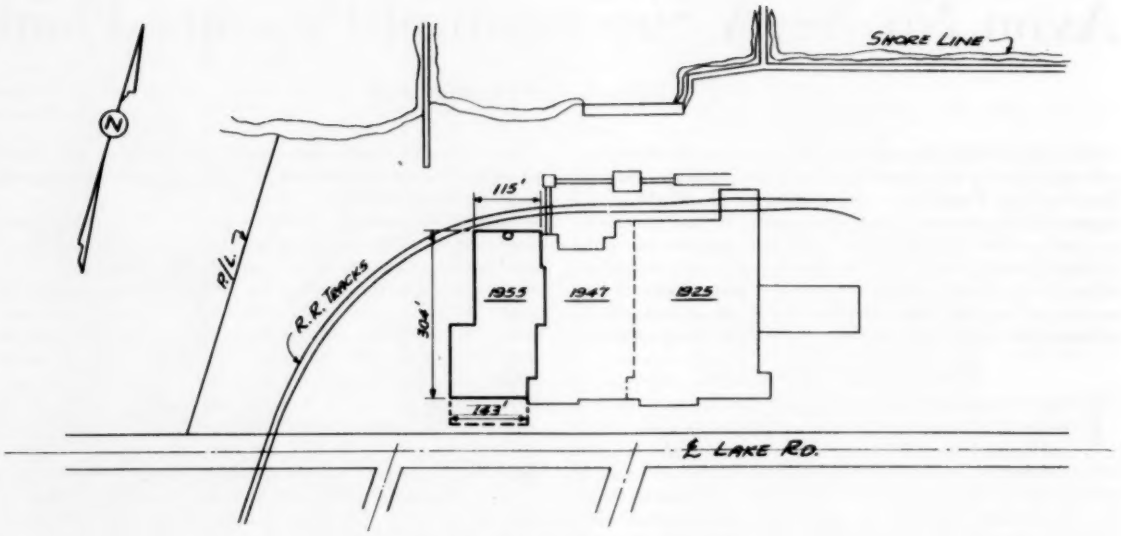


FIG. 1 PLOT PLAN OF AVON ADDITION

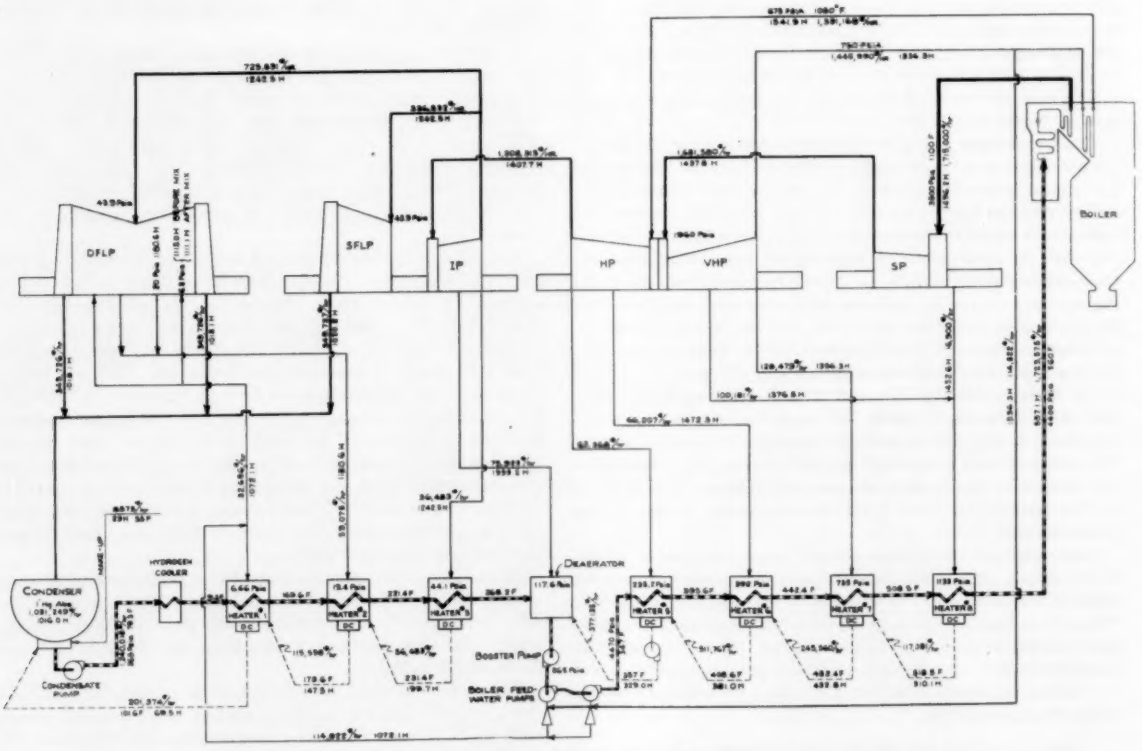


FIG. 2 SIMPLIFIED HEAT-BALANCE DIAGRAM

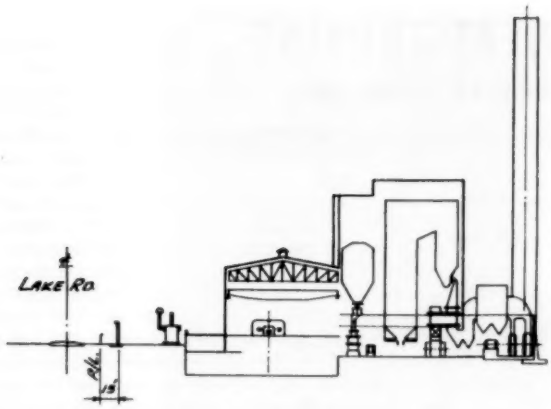


FIG. 3 PLANT CROSS SECTION

signed to accommodate two units and will be located directly between the turbines and boilers on the turbine operation floor level. Main electrical controls will be located in an existing control room installed at the time of the construction of the original section of the station.

Fig. 2 shows a simplified heat-balance diagram for this unit. The estimated net plant heat rate at rated load and back pressure is 8620 Btu per kwhr with two steam-turbine-driven boiler feed pumps in service. With one turbine-driven and one motor-driven feed pump in service, the plant net heat rate will be reduced to 8605 Btu per kwhr.

STEAM GENERATOR

Steam will be supplied to the 250,000-kw supercritical-pressure turbine by a C-E Sulzer monotube steam generator as shown in Fig. 4. It will be a pulverized-coal-fired unit of the dual-furnace design as manufactured by Combustion Engineering, Inc. Corner firing using tilting tangential burners will be employed to burn low ash-fusion-temperature coal. The unit will be designed to deliver 1,715,000 lb of steam per hr to the turbine throttle at 3500 psig and 1100 F with single reheat to 1050 F. The boiler has no drum and is of the forced-circulation once-through type. In many respects it is similar to Sulzer monotube units in operation in European power plants. The main boiler feed pumps furnish sufficient head to force the boiler water through the continuous-tube circuits of the various sections of the boiler as well as the feedwater and main steam lines.

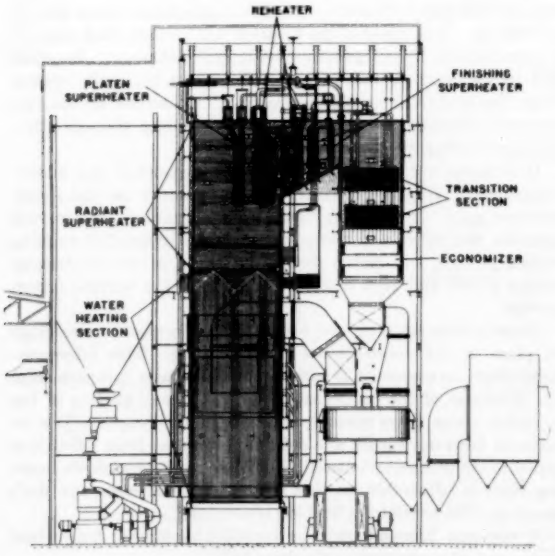


FIG. 4 BOILER CROSS SECTION

After leaving the two parallel lines of feedwater heaters, the feedwater is divided into a total of four circuits. Two feedwater circuits supply an economizer for each furnace. The feedwater flow to the economizers is controlled in each circuit by individual feedwater control valves. The feedwater then flows to the water walls which occupy approximately the bottom half of the wall surface in each furnace. No conventional downcomers outside of the boiler are required to supply the water-wall section. The boiler water makes a number of up-and-down passes in the furnace water-wall section before entering the horizontal-tube transition zone which is located in a relatively cool convection region. The transition zone is where water is converted into steam. As will be discussed later, very careful control of feedwater quality is maintained to prevent the deposition of solids in the transition zone and superheater.

On leaving the transition zone, each of the two steam circuits per furnace is subdivided into two parallel circuits through a series of radiant, platen, and finishing superheaters. These eight steam circuits include appropriate stop and bypass valves, which are used during start-up and abnormal operating periods. To maintain uniform steam-temperature distribution to the turbine, the leads from the stop and bypass valves are connected into a mixing header.

In developing a logical and balanced design, it will be necessary to obtain variations from some of the arbitrary rules in today's Boiler Codes. This results from the fact that the codes as written did not have this type of boiler in view. All the variations have been developed and approved by the Industrial Commission of Ohio. As an example, the safety-valve capacity cannot be located on the drum as there is no drum. Therefore, it is intended that all safety-valve capacity will be located at the superheater outlet and the safety valves will be set above the operating pressure, so that they will not affect operation unless the automatic control and bypass valves do not function. Thus, safety valves are even more in the category of occasional departures than is the normal case on utility boilers.

TURBINE-GENERATOR

The Westinghouse turbine-generator, a cross section of which is shown in Fig. 5, is designed for a maximum throttle flow of 1,715,000 lb of steam per hr at a pressure of 3500 psig and temperatures of 1100 F with reheat to 1050 F. With two turbine-driven boiler feed pumps in service, the maximum capability of the main turbine is 250,000 kw. The 3600-rpm inner-cooled generator has a rating of 273,460 kva at 22 kv, 0.85 power factor, 0.64 short-circuit ratio, and 45-psig hydrogen pressure. Provision is made to operate the generator at 60 psig. A 1200-kw exciter driven by a 715-rpm induction motor operating at 375 volts supplies excitation current.

The tandem-compound, triple-flow turbine receives supercritical-pressure steam which at full load is expanded to 2000 psia in the steam-cooled superpressure section. The steam then passes to the very-high-pressure section where its pressure is reduced to 750 psia. Reheated steam returns to the high-pressure turbine at 1050 F where it expands to 240 psia. After passing through the intermediate-pressure turbine, the steam enters the triple-flow low-pressure sections at 45 psia. The design condenser pressure is 1.0 in. Hg.

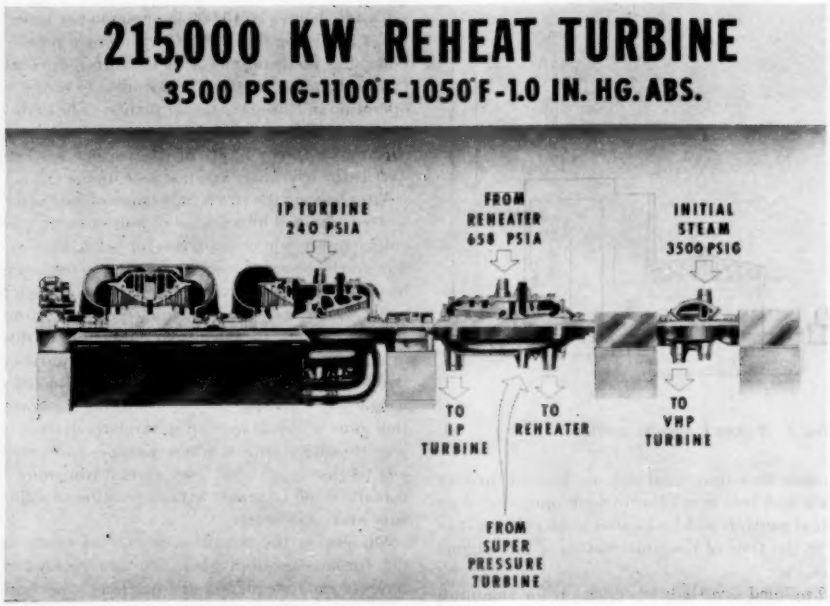


FIG. 5 TURBINE CROSS SECTION

The entire control and protective system on this turbine is hydraulically operated. The main steam controls include servomotors for moving the governor valves which determine the flow to the turbine, the speed-responsive governor, and the speed changer, which adjusts the speed range in which the governor will control the servomotors. The interceptor valves are closed during start-up and up to a steam flow of 30 per cent of maximum.

The protective system includes the main steam throttle valves, the reheat stop valves, the overspeed trip valve and its connected devices for tripping the main steam throttle, the governor, interceptor, and reheat stop valves whenever the turbine reaches a predetermined overspeed. A load-limit device also is provided to limit the maximum opening of the governing valves to any chosen amount. In addition, the turbine is protected against low vacuum and low oil pressure. The throttle and reheat stop valves have only two positions; namely, closed or fully open.

The governing valves function in parallel to allow full peripheral admission of steam to the first stage of the turbine. As a result, therefore, the throttling loss at partial loads will be greater than for machines where the governing valves open sequentially.

Although the final complement of supervisory instruments has not yet been decided upon, it is likely that several new ones will be added to those normally utilized on subcritical machines. In accordance with current practice on the system, provision is being made for remote control of the turbine-generator.

By maintaining the same turbine-room width for the station extension, it will be possible to use the existing 125-ton crane in the new section. An additional 175-ton crane will be installed with the new turbine. Current plans are to have the generator shipped as a complete unit weighing approximately 625,000 lb. Both the existing and the new crane will be used to transfer it to the turbine-generator supporting foundation.

BOILER FEED PUMPS

There will be three six-stage main boiler feed pumps each capable of delivering 2285 gpm (1,026,000 lb per hr) at a total head of 11,580 ft. Two of the pumps will be driven by steam turbines and the third by two 4000-hp motors in tandem.

The two boiler-feed-pump-drive turbines will receive steam from the "cold reheat" line and will deliver exhaust steam to the lowest stage extraction line. Part or all of the steam will flow to the lowest stage extraction heater to supply its requirement. The balance, if any, will return to the turbine in a reverse direction from the conventional extraction flow, and continue through to the main condenser.

The maximum speed of the boiler feed pumps and their drive turbines, as well as the output speed of the speed-increase gear, will be 7575 rpm. This corresponds to a maximum power output of 7650 hp. The speed of the turbines will be controlled through a conventional flyball governor with speed changer. Impulses will be received by the governor through a hydraulic system from the feedwater regulating valve. The output of the two motors in tandem will be delivered to the pump through a hydraulic-coupling speed-increase gear.

It is tentatively planned to control the speed of the motor-driven pump from a governing device mounted on the speed-increase gear. The output signal from a hydraulic system will position the speed-regulating tube of the hydraulic coupling which, in turn, will match the output speed of the electrically driven pump and that of either or both of the turbine-driven pumps.

These pumps are equipped with labyrinth-breakdown bushings in place of the conventional shaft packing. This labyrinth-breakdown arrangement is based upon cold-water injection from the discharge of the condensate pump. A small portion of the injection water flows inwardly into the pump proper. The remainder flows outwardly to a collection chamber from which it is piped to the condenser hotwell. The labyrinth-breakdown bushing itself is adjustable, so that concentricity between the shaft sleeve and the breakdown bushing is assured.

A common baseplate will be installed under the boiler feed pump and all equipment in the drive train.

The purpose of the dual-motor drive is to limit the starting cur-

rent so that standard 4000-volt switchgear can be used. In starting, power is applied to one motor until up to full speed.

Choice of Pump Arrangement. In establishing the boiler-feed-pump arrangement, it was first decided that such pumping would be accomplished in one stage rather than two. With a maximum feed-pump discharge pressure of 4730 psig, it was found that the savings in pump power in handling relatively cool water (339 F) exceeded the savings which would have resulted from designing the high-pressure feedwater heaters for a lower pressure. The use of such lower pressures for heater design would have been permitted by the addition of a second stage of feedwater pumping which would have supplied something in the order of 50 per cent of the pumping requirement after the feedwater had left the highest pressure feedwater heater. The added cost of two sets of pumps, as compared with one, was another factor pointing toward a single stage of pumping. As a result, all feedwater pumping will be accomplished in one stage ahead of the high-pressure heaters in the conventional manner. To assure adequate NPSH to the boiler feed pumps at all times, low-power booster pumps are arranged to receive feedwater from the deaerator and deliver it to the main boiler feed pumps.

The equipment and controls are being designed so that any two or three of the pumps can operate in parallel satisfactorily. Present expectations are that the two turbine-driven pumps will be capable of operating in parallel down to 30 per cent of full load.

Feedwater Heaters

The four high-pressure feedwater heaters are designed for a water-side pressure of 4730 psig. Since these four heaters are located on the discharge side of the boiler feed pumps, it has been necessary to establish their operating pressure to meet the pump-discharge pressure required to deliver 3500 psig steam to the turbine throttle. The three low-pressure condensate heaters are designed for 300 psig, the maximum discharge pressure of the condensate pumps. The deaerating feedwater heater is designed for 125 psig. In addition to the feedwater heaters, a hydrogen cooler, condensate cooler, and gland-steam cooler are located in the condensate cycle. A steam regulator controls the flow of steam to the gland cooler. A diagram of the condensate and boiler feedwater cycle is shown in Fig. 6.

After study of the cost factors involved, it was decided to employ two heaters in parallel for each of the high-pressure feedwater-extraction stages. Cost factors considered in this study included cost of the heat exchanger, piping cost, valve cost, and space evaluation for each of the arrangements. With two strings of heaters, it was deemed unnecessary to design piping for bypassing an individual heater. As a result, difference in valving costs assumed significant proportions. Should any high-pressure heater require maintenance, that entire string of heaters will be taken out of service. It is hoped that it will be possible to carry almost full load under this condition.

The high-pressure heaters are of the U-bend type with integral desuperheaters and drain coolers. Each of the drain coolers is designed for a 10-F approach to the temperature of the feedwater entering the heater. Tubes are of 15 BWG minimum wall monel, welded to the tube sheet.

The low-pressure heaters are the conventional U-tube design including integral desuperheaters and drain coolers. The second lowest pressure heater will be located in the condenser neck. All units are designed with integral tube sheets and channels with separate removable channel covers. As in the case of the high-pressure heaters, all low-pressure heaters are equipped with 10-F approach drain coolers. Tubes in the low-pressure heaters will be of 18 BWG admiralty metal.

Original plans had called for the drains of the lowest pressure heater to be pumped into the condensate stream between the first and second lowest pressure heaters in order to obtain the best possible heat rate. After consideration of water-conditioning problems, however, it was decided to send these drains directly to the condenser. The possibility of significant copper and iron pickup by the extraction drains in cascading through the low-

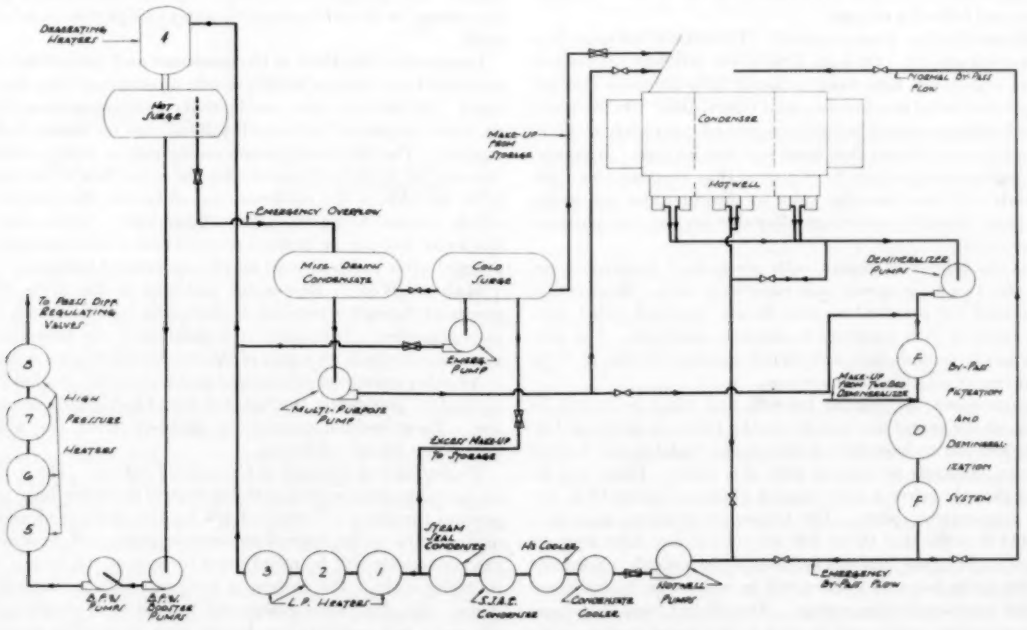


Fig. 6 Condensate and Boiler-Feedwater Cycle

pressure heaters made it desirable to demineralize low-pressure drains as soon as possible.

The deaerating feedwater heater consists of two vertical steel sections equipped with spring-loaded spray valves and deaerating element. These deaerating heater sections, with internal direct-contact vent condensers, are mounted on, and supported by, a horizontal cylindrical steel storage tank of 6200 cu ft capacity.

Water-Quality Control

The once-through supercritical-pressure boiler design imposes stringent limitations on total solids in the boiler feedwater. Solids formerly blown from the boiler drum can lead to deposit formation in the unit if not removed by other means. Solids in the steam may deposit on the turbine blades.

In the latest 2400-psig controlled-circulation boiler at Eastlake Station, the feedwater pH is maintained at 9.0 to 9.2 by adding morpholine or ammonia. Feedwater conductivity (undegassed) of 1.0 to 1.5 micromhos is satisfactory. In the supercritical-pressure cycle, the feedwater will have the following limitations:

pH 9.0 – 9.5	by adding ammonia or hydrazine
Total solids	50 ppb maximum at condensate pump discharge
Total iron	10 ppb maximum
Total copper	10 ppb maximum
Silica	20 ppb maximum
Oxygen	7 ppb maximum

A condensate-bypass filtration and demineralization system will be used to maintain the solids within the required limits. The bypass filtration and demineralization system serves three major purposes: namely, (1) protects feedwater cycle from condenser-tube-sheet weepage (leakage between tube and tube sheet) and condenser-tube rupture leakage, (2) "polishes" make-up water from a two-bed demineralizer system, and (3) removes corrosion products from the cycle.

Two secondary purposes of the bypass filtration and demineralization system are (1) to polish miscellaneous drains before returning to the cycle, and (2) to clean up the cycle on original start-up and following outages.

Condenser-Leakage Contamination. Tube-sheet weepage is a controversial subject. On Lake Erie water excellent results and long life expectancy have been achieved with red-brass and admiralty tubes rolled into muntz-metal tube sheets. On the latest 1800 to 2400-psig units tube tightness proved completely satisfactory and it was believed that there was zero leakage. However, recent highly sensitive tests have proved that weepage does exist. The solids introduced into the water by this condition were found to be more than the maximum allowable for the supercritical-pressure cycle.

A double tube-sheet design with pressurized condensate between the two tube sheets was considered first. However, it was decided for practical reasons to use standard rolled tube sheets sprayed with neoprene to promote tightness. The condenser and hotwell designs will permit demineralization of 95 to 100 per cent of all tube-sheet weepage.

The condenser, rectangular hotwell, and water boxes will be divided on the center line parallel to the tubes as shown in Fig. 7. In addition to loose-fitting tube-support plates, the hotwell will have divisions transverse with the tubes. These will be located directly below a tube-support plate and about 18 in. below a tube-support plate. The transverse divisions have been provided in order that 95 to 100 per cent of any tube weepage will be caught in the four end sections of the hotwell. Normally all water in the four end sections will be sent to the bypass filtration and demineralization system. The effluent from the bypass demineralization system will be sent to the center section of the condenser for redeaeration by the condenser before the water is returned to the cycle. The full-load design flow for the bypass system is 625,000 lb per hr or 50 per cent of condenser flow, including low-pressure heater drains.

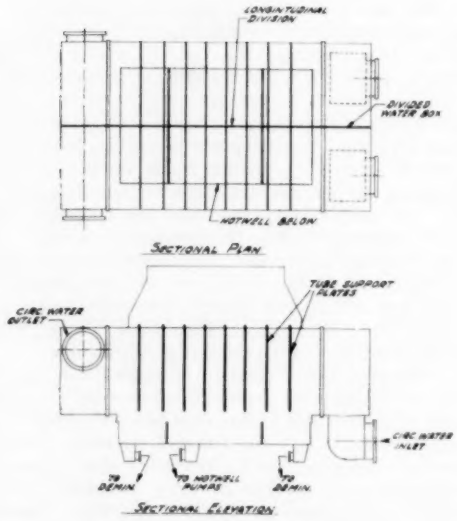


Fig. 7 Condenser Arrangement

The bypass filtration and demineralization system has been designed for 100 per cent condenser flow during emergency operation. Tube rupture will be detected at the hotwell pump by precise conductivity measurements. When abnormally high conductivity is detected, all condensate will be sent through the bypass filtration and demineralization system. By designing the bypass system for 100 per cent condenser flow, the unit can hold full load for 24 hr or longer with no danger of cycle contamination. By comparison to other supercritical-pressure plants, the amount of stored high-purity water (358,000 lb) is relatively small.

Longitudinal divisions of the condenser and hotwell have been provided to determine on which side a condenser tube has ruptured. In this case, also, conductivity measurements on each of six outlet nozzles of the hotwell will indicate the location of tube rupture. The circulating pump on the side in which failure occurs can be turned off, permitting the water box to be drained. With one side of the condenser out of service the turbine load will be reduced to approximately 70 per cent. At the same time the water box can be drained and the end of the ruptured tube plugged, after which the unit may be operated at full load.

Make-up Water. Raw-water make-up to the cycle will be processed through a two-bed demineralizer equipped with a vacuum degasifier. This make-up is polished in the mixed-bed demineralizers which are a part of the condensate-bypass system.

All other sources of condensate make-up at the station can be utilized by processing through the mixed-bed bypass demineralizer. These include evaporated make-up from the adjacent boilers and stored condensate.

Prevention and Removal of Corrosion Products. The cycle has no special materials, such as stainless-steel or rubber-lined pipe to prevent corrosion. Control of pH by the addition of ammonia or hydrazine at the hotwell condensate pump and deaeration of the condensate will be relied upon to protect the metals in the preboiler cycle. The system is equipped with a standard condenser with deaeration guaranteed to 0.03 cc of oxygen per liter (43 ppb) and a deaerating heater with deaeration guaranteed to 0.005 cc of oxygen per liter (7 ppb). The high-pressure

heater drains are returned to the deaerating heater. The low-pressure heater drains will be returned to the end sections of the hotwell for bypass polishing.

The filters ahead of the bypass-polish demineralizers will be relied upon to remove the corrosion products in the condensate. The nature of the filter material has not been established but work is in progress to determine experimentally the most desirable filter material. Postdemineralizer filters will be used to remove traces of resin and elutriated materials from the ion-exchange beds.

Power Piping

The main steam piping for this unit shown diagrammatically in Fig. 8, presented problems not previously encountered on conventional subcritical-pressure units. Eight main steam leads of 6⅝-in. OD × 4-in. ID (ASTM Specification A-376, Grade TP316, austenitic material) were established from the superheater outlets to the boiler stop and bypass valves. Each of these leads is approximately 300 ft long. From the boiler stop and bypass valves, eight austenitic 6⅝-OD × 4-in. ID steam leads will run to a 16-OD × 10-in. ID austenitic mixing header. From the mixing header, four austenitic 7⅜-in. OD × 5-in. ID turbine leads will run to a transition section of austenitic to ferritic material. From the transition sections, the four turbine

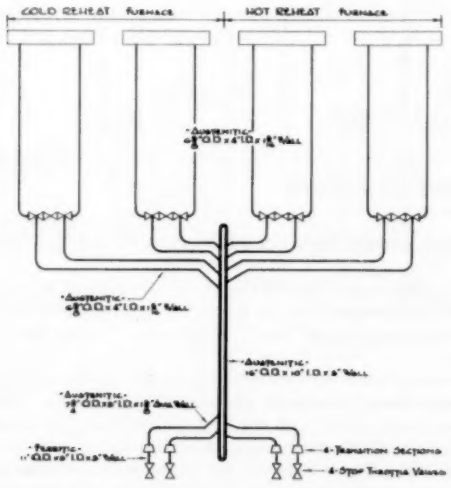


Fig. 8 Diagram of Main Steam Piping

leads to the turbine stop valves will be ferritic material (2¼ Cr, 1 Mo) ASTM Spec. A-335, Grade P22, 11-in. OD × 5-in. ID.

Owing to the operating conditions of 1100 F at the turbine throttle, it was possible to have a metallurgical choice between ferritic or austenitic alloys. Tests were performed on Types 316 and 347 in coarse and fine-grain specimens by heating a 20-in. length of small-diameter heavy-wall pipe to 1440 F, clamping, and then cooling to 1200 F to obtain plastic strain. The conclusions of these tests indicated that the hot ductility of fine-grain Type 316 is much better than coarse-grain materials. These facts led to the decision to specify relatively thin-wall fine-grain austenitic pipe for the eight main steam boiler leads. It is proposed that the required pipe will be manufactured by the extrusion method of a fine-grain material with the same manufacturing tolerances as "seamless" for which the wall-thickness variation is plus or minus 12½ per cent. The normal 0.065-in. corrosion allowance was included in the pipe-wall calculations. The lengths available from the extrusion press in the 4-in-ID size up to 1½-in. wall are approximately 30 ft.

By procuring long lengths the welded joints will be kept to a minimum. Root-pass welds in the main steam lines will be made using the inert-gas tungsten-arc method combined with the use of consumable backing rings. This procedure utilizes a shielded arc to draw an inert-gas-backed ring into the joint to form a first-pass bead completely fused with the parent metal and without any crevices. The welds are to be made without preheating. All the welds will be ground flush and inspected by radiography and fluid-penetrant methods.

Design pressures and stress allowances were determined for the main steam piping from the superheater outlets to the boiler stop and bypass valves (under jurisdiction of the Boiler Code) by taking the turbine throttle pressure of 3500 psig, 1100 F and adding a design allowance of 5 per cent plus the calculated pressure drop in the piping, valves, and mixing header back to the superheater-outlet headers from the turbine stop valves.

The resulting superheater-outlet-header design pressure was approximately 3980 psig at 1110 F and all pipe-wall calculations for the austenitic materials to the stop and bypass valves were determined on this basis. The stress allowance of Grade TP316 was interpolated for 1110 F temperature.

The low velocity of 9100 fpm through the main steam leads became necessary to minimize an already large pressure drop in the piping and valves of approximately 275 psig. From the boiler stop valves through the mixing header to the turbine stop valves the austenitic and ferritic main steam leads were designed under ASA B31.1 Piping Code and varying design pressures with 1100 F temperature were used.

The boiler-feedwater piping was designed by utilizing the higher allowable stress present in the new ASTM Spec. A-106, Grade "C." Owing to reluctance of domestic mills to roll Grade "C" carbon-steel pipe, the order was placed in Germany.

Boiler-Control Systems

The once-through boiler has five control systems:

1 Feedwater regulation.
2 Main steam temperature.
3 Turbine-bypass system.
4 Combustion control.
5 Reheat temperatures.

In terms of conventional American boiler practice, the turbine-bypass control system is the most unusual feature necessitated by the once-through design. The feedwater regulation is based on temperature rather than on drum level.

Hydraulically operated controls will be employed for the feedwater, main steam temperature, and turbine-bypass system as shown diagrammatically in Fig. 9. These will be similar to those with which a great deal of experience has been had in subcritical once-through boilers in Europe.

Combustion control to adjust fuel and air supply in accordance to load will be of electropneumatic design. Many new features, such as matching firing rate to generator load and air flow corrected by oxygen content of flue gas, are being contemplated.

Reheat steam-temperature control will be of electrical design. Since reheater surfaces exist in each of the two furnaces, the temperature control will have two stages. Burner-tilt temperature control will be employed in each furnace. The final reheat temperature will employ sprays for desuperheating as backup protection.

A brief description of the boiler heating surfaces is necessary to understand the basic controls. The boiler has four separate circuits, each consisting of an economizer, water walls, transition section, radiant superheater, platen superheater, and finishing

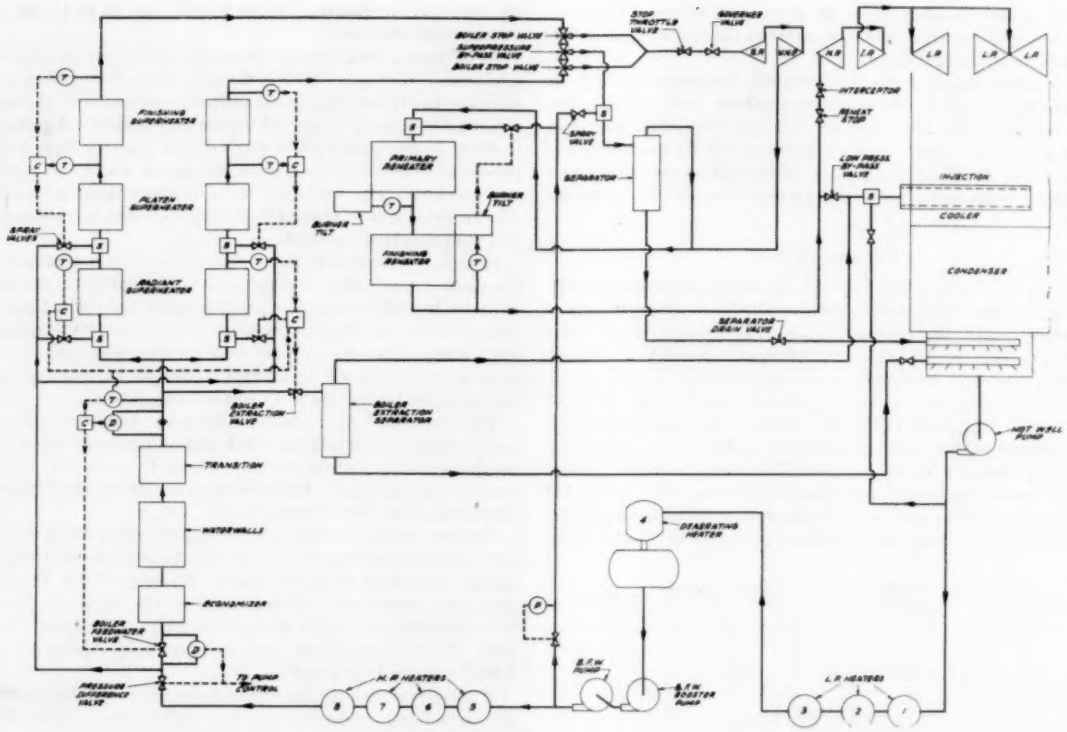


FIG. 9 BOILER-CONTROL AND TURBINE-BYPASS CYCLE

superheater. The individual parallel circuits extend from the pressure-difference regulating valves to the mixing header just before the turbine throttle. Each circuit has independently controlled valves to regulate feedwater flow, boiler extraction, spray water for steam-temperature control, and flow to turbine or flow around turbine to the reheater.

Feedwater flow in each of the circuits is controlled by the steam-flow nozzle and the temperature leaving the transition zone. The steam-flow nozzle is used as an anticipation device and the final control is provided by temperature. Maintaining the proper temperature at the outlet of the transition zone matches feedwater flow to firing rate.

The pressure-difference regulating valve is used to maintain an approximate 30-psi drop across the feedwater regulating valve. The use of the pressure-difference valve compensates for variation in feed-pump pressure and boiler pressure so that the feedwater flow depends exclusively on the position of the feedwater valve.

The speed of the feed pumps is regulated by the position of pressure-difference valves. If the flows are slightly unbalanced in the four circuits, one of the pressure-difference valves will be opened slightly more than the others. The pump speed will correspond to the pressure-difference valve open the widest. In this manner the pumps will be running fast enough to maintain proper flow in each of the four circuits.

The main steam temperature in each circuit is regulated in two stages by means of spray-type desuperheaters located at the inlet of the radiant superheater and at the inlet of the platen superheater. Each of the two steam-temperature controls functions essentially alike.

The spray at the inlet of the radiant superheater is controlled by the temperature leaving the radiant superheater with anticipating impulse from the temperature leaving the transition zone.

The spray at the inlet of the platen superheater is controlled by the temperature leaving the finishing superheater with anticipating impulse from the temperature leaving the platen superheater.

The spray water for each circuit is taken from the line between the feedwater valve and the pressure-difference valve. The flow of spray water varies with the flow of feedwater prior to adjustment by the two-element temperature control.

Provision has been made for extracting water from the system at the outlet of the transition section, the purpose being to reduce the flow through the superheater during unusual operating conditions. The three primary functions of the boiler-extraction system, which is controlled by the same two-element temperature control as the spray valve at the inlet of the radiant superheater, will be to operate during loss of fire, under conditions of extremely low load, or during a hot restart. The use of this extraction system will limit sudden temperature reduction in the high-temperature superheater and piping on loss of fire, or make it possible to maintain proper temperature suitable for the turbine during hot restart.

The boiler-extraction valves, which are located in each circuit after the transition zone, discharge to the condenser and will be locked closed on a cold start. Interlocks prevent the valves from opening if the condenser cannot condense the steam without excessive temperatures.

The turbine-bypass system is required for cold starts, hot restarts, turbine trips, and abnormal operating conditions. Flow of steam or water bypasses the turbine through the superpressure bypass valve in each circuit as shown in Fig. 9. The superpres-

sure bypass valves discharge into a separator. The water from the separator is discharged into the condenser. The steam from the separator discharges into the cold reheat lines. The steam travels through the reheater out the low-pressure bypass valves on the hot reheat lines. The steam from the low-pressure bypass valves discharges into the injection cooler in the neck of the condenser. The steam at the superpressure-bypass-valve outlet is sprayed to keep the temperature equal to cold reheat temperature. The steam in the injection cooler is sprayed to limit the turbine-exhaust-hood temperature.

The primary function of the boiler stop valves in each circuit is to prevent water from entering the turbine. On cold start-up, water is pumped through the entire boiler to the condenser through the separator. Only when proper steam conditions are established will steam be admitted to the turbine.

The combination of the superpressure bypass valves and boiler stop valves also prevents steam from entering the turbine at temperatures excessively above or below those allowable at any given time. The bypass valve also acts as a relief valve which will open at 105 per cent pressure.

Station Heating and Ventilating

High-Temperature Water-Heating System. The conventional low-pressure saturated steam for coil and general building heating could not be used economically because of the absence of saturated steam from a boiler drum and the desirability of minimizing make-up of high-purity feedwater required for the boilers. Therefore, a high-temperature, closed-cycle forced-circulating hot-water system was selected as shown diagrammatically in Fig. 10. It permits greater flexibility in pipe runs, smaller size of pipe, and the elimination of trap maintenance.

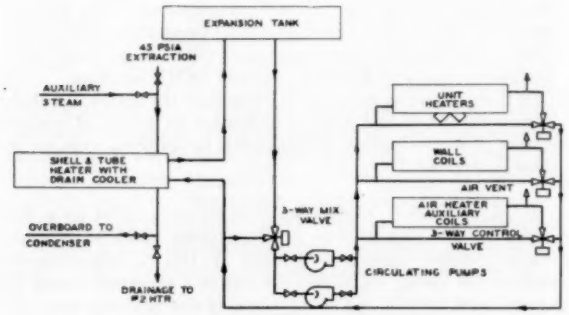


Fig. 10 High-Temperature Water-Heating System

To obtain high-temperature water, superheated steam will be extracted from the cycle at the 45-psia stage and passed to an ordinary feedwater heater-type heat exchanger equipped with a drain cooler. The condensate will then flash to the No. 2 low-pressure heater. In the high-temperature water circuit, a pump will continuously circulate water at approximately a constant flow rate through the heat exchanger and on through the air-heating coils and unit heaters. In general, control will be provided by a thermostatically controlled bypass valve to bypass water around the heat exchanger. A spare pump will be provided.

For starting up unit No. 8 and for heating the building during loads 30 per cent or lower or shutdown periods, steam will be provided via a pressure-reducing station from the boiler drums of unit no. 6 or no. 7.

Combustion-Air Preheating. The current practice of reducing building cubage to a minimum and the constant effort to increase plant efficiencies requires a carefully engineered method of introducing combustion air into the building.

The combustion air will be drawn through wall louvers and heating coils at a low elevation in the precipitator bay. The air rises to the top of the boiler house absorbing approximately 75 per cent of the boiler and equipment radiation losses which amount to approximately 1.5 per cent of the boiler input. From the top of the boiler house, the air will be drawn down through large air shafts which are connected by ducts to the forced-draft fans.

A portion of the combustion air will be drawn through a 30-ft-deep air well, through heating coils by a large centrifugal fan and discharged into the turbine-room basement near the feedwater heaters.

The balance of the building heating will be supplied by thermostatically controlled unit heaters.

Auxiliary heating coils will be installed in the air-inlet side of the air heaters to maintain proper air-inlet temperature during start-up and during low load.

Air Conditioning and Ventilation. The boiler control room, the relay room, the electrical control room, and the laboratory will be air conditioned. Fresh air will be taken from outside the plant and a positive pressure will be maintained in the control rooms to permit occupancy in case of fire in the plant.

Roof ventilators, gravity type, will be installed on both the turbine house and boiler house. The turbine-house ventilators will be motor operated and manually controlled. The boiler-house ventilators will be opened automatically whenever the air-shaft temperatures exceed 135 F. Both sets of ventilators will have fusible links with weighted dampers set to open in case of fire, permitting positive escape of high-temperature gases and smoke and preventing failure of the roof structure.

Stack

One of the problems in the existing plant has been air pollution. In order to determine how this addition could be constructed to be a minimum nuisance, it was decided to test a model of the existing plant, together with one of the proposed new addition, in the New York University wind tunnel. After trying various locations for the new stack and various stack heights, finally it was determined that the best location would be directly back of the boiler house with a height of 400 ft above grade.

Several studies were made to determine of what materials the stack should be constructed. In the past, extensive use had been made of steel stacks with brick linings. However, when it was decided to go above 300 ft, it was found that reinforced-concrete stacks with brick linings were more economical.

The substructure for the stack is a 42-ft octagonal concrete pad of maximum depth of 6 ft. The stack tapers from 28-ft OD at the base, to 18-ft 6-in. OD at the 400-ft level. Thickness of concrete varies from 1 ft 9 in. at the base to 6 in. at the top. The stack will be lined with 4 in. of dense brick and will have 1-in. fiberglass insulation to separate the brick from the concrete.

Job Status

Engineering work on the supercritical-pressure addition to Avon Station began in August, 1955. As was mentioned earlier, the entire design had not been established completely at the time of the writing of this paper. Some studies are still being made, and therefore it is entirely possible that changes may be made in some of the design information that has been presented. Throughout the paper emphasis has been placed on departures from design occasioned by the use of supercritical pressure in a large central-station single boiler-turbine-generator unit.

Construction work has begun already, and some of the footings are in place. Present plans are to have the unit in operation late in 1958.

Conclusions

This plant should contribute considerably to the advancement

of the art of power generation. It is a third approach to the use of supercritical pressure in large central stations. Many of the features described in the paper have not been incorporated heretofore in any other project. Although it does not have the best projected heat rate, it appears at present to be the most logical selection of steam conditions in the supercritical-pressure range. This is evidenced by the fact that others are now contemplating units operating in the 3500-psig range.

ACKNOWLEDGMENTS

The author's thanks are due to his colleagues in the Civil and Mechanical Engineering Division, principally Messrs. H. L. Colgin, R. L. Dombey, N. F. Gill, R. W. Ott, S. B. Rock, R. B. Shumaker, C. D. Trump, and G. H. Walker who contributed information for the final manuscript, and to the several manufacturers who co-operated in the preparation of the paper.

Discussion

J. L. ALLEN.[2] In presenting this discussion, the intent is not to be critical of the paper itself but rather to present, for consideration, a comparison of several systems described by the author for Avon No. 8 unit with like systems being employed at Eddystone unit No. 1, the first supercritical-pressure unit of the Philadelphia Electric Company.

As presented, the paper of course is limited somewhat in scope, and many phases remain to be discussed. The many questions remaining unanswered at this time undoubtedly will be the subject for discussion in subsequent papers.

The Eddystone unit is a cross-compound machine rated at 325,000 kw and, except for the steam condition of 5000 psi, 1200, 1050, 1050 F (two stages of reheat and nine feedwater heaters) the heat-cycle arrangement is somewhat similar to that of Avon No. 8.

Two lines of heaters are employed with crossties only at the condensate-pump discharge, the outlet of the cycle, and at the deaerators, thus minimizing the valving required and providing complete flexibility for bypassing operation in the feedwater circuits.

The pumping system will employ three condensate pumps discharging through two lines of three low-pressure heaters into the deaerators. Two low-pressure boiler feed pumps, with double-suction first stages (thus reducing to a minimum the NPSH requirement and avoiding the use of a booster pump), take suction directly from each deaerator and pump through five stages of moderately high-pressure heaters; i.e., of 3000-psi construction.

Here then two differences of opinion exist. Heaters of 3000 lb were selected because of several factors, among which were considered: (1) It was thought good judgment to continue using rolled tube joints at this pressure because of excellent experience with several other units now operating on the Philadelphia Electric Company system under somewhat the same conditions. (2) Unlike Avon No. 8, the savings in heater cost and the effects of heat input by the intermediate and high-pressure pumps on the boiler design provided an economical justification for placing them after the final stage of feedwater heating.

The feed pumps thus far identified as low, intermediate, and high are three pumps in series; the first two being 3600-rpm direct-connected motor-driven units with the high-pressure pump being turbine driven. The problem of starting and light-load operation is of paramount concern with the Eddystone unit as well as with Avon No. 8 or any other supercritical-pressure units for that matter.

For this reason, by allowing the low and intermediate boiler feeders to develop enough pressure to carry approximately 30 per cent load above critical pressure, the high pressure (i.e., the turbine-driven pump) would be pumped through during this phase of operation. From 30 to 100 per cent load, the turbine-driven pumps would act as variable-speed units receiving control signals in the same manner as that used to control the Avon No. 8 variable-speed machines.

By the arrangement thus described it is hoped that the complications and difficulties associated with speed-increasing gears, hydraulic couplings, and shifting of controls will be avoided.

As on Avon No. 8, water treatment and tightness of the condenser are of great concern. It is believed that we have carried the solution of these problems two steps further than indicated at Avon: (1) The condenser hotwell is so divided that either end or the center may be pumped out to the demineralizing system, thus guarding not only against tube weeping at the tube sheets but also possible damage caused by vibration or a split tube in any section of the condenser. (2) All make-up (or a predetermined amount of continuously recycled condensate) is injected into the heat cycle after the condensate-pump discharge rather than into the condenser, thus avoiding possible contamination of the newly demineralized condensate or of dilution of the hotwell effluent.

The boiler and its controls are essentially the same as described by the author for the Avon unit.

It should be understood that many more comparisons could be given between the two plants. Time however does not permit going deeper into the many facets of the projects, such as pump sealing and leak-off, soot blowing, preheating of boiler combustion air, stack-gas cooling, and turbine-bypass systems. Many innovations to normal engineering practices are becoming a reality with these and other superpressure units and will be subjects for many informative hours of listening at future ASME meetings.

Of this you can be sure, and to paraphrase the author, "A challenge is being met and a real contribution to the art of central station power generation is being made by all those engaged in these enterprises."

E. M. POWELL.[3] The author is to be commended for a fine paper and his management for the policies which have made this project and presentation possible. In this day of heavy competition for young engineers, the author's expressed intent to attract their interest and convey some of the challenge that still faces engineers in the power-generation industry is particularly significant.

The willingness to select steam conditions which at the moment appeared to be one step in advance of the most economical unit as a contribution to the advancement of the art is an excellent indication of the progressive leadership of The Cleveland Electric management. It is a privilege to be associated with them on this project.

In most respects the steam-generating equipment follows the standards of design as developed by the writer's company. Superficially at least, it resembles the No. 4 unit at Eastlake Station. The most notable exceptions result from the characteristics of steam at supercritical pressure—the inability to separate water from steam following partial evaporation and the continual increase in temperature as the fluid flows through the various circuits from the economizer inlet to the superheater outlet.

Not only does this require a revised approach to feedwater purification but to the operating procedures and in the design of the control system. In this design, particular emphasis has been placed on the control of temperature in all parts of the unit as outlined in the paper.

Some years ago in thinking ahead to the day when units de-

[2] Engineer, Philadelphia Electric Company, Philadelphia, Pa. Assoc. Mem. ASME.

[3] Assistant Chief Engineer, Combustion Engineering, Inc., New York, N. Y. Mem. ASME.

signed to operate at pressures above the critical would become a reality, our management became convinced that one of the most essential items would be a control system carefully integrated with the design of the steam-generating equipment. This led to a license agreement with Sulzer Brothers of Switzerland who had developed such a control system which could be adapted readily to suit our needs. That firm also had an extensive background of experience in the successive operation of once-through steam-generating units.

While there has been nothing in Europe which exactly duplicates the conditions for Avon No. 8, there has been considerable experience with advanced steam cycles and once-through boilers. The basic principles and operating procedures are so similar as to require moderate extrapolation. A brief review of some of the high lights of European experience might be interesting.

Sulzer Brothers contracted for its first monotube steam-generating unit in 1929. Since then the company has built and placed in service nearly 90 units covering a wide range of operating conditions. The maximum pressure at the superheater outlet of any one of the installations is 2062 psig. Similarly the maximum steam temperature developed is 1112 F and the largest capacity is listed at 660,000 lb of steam per hr. Some of these units are equipped with reheaters and one such operates at 1015 F.

Many of the 90 monotube units have been installed in central stations but the majority serve in industrial plants as topping units with the high-pressure turbine exhaust going either to process or to low-pressure turbines.

These industrial installations in Europe, however, do not operate as separate entities. To appreciate the significance of this one must be familiar with the degree to which the government in many European countries becomes involved in the operation of industrial as well as power projects. This is accomplished in the name of conservation of resources such as fuel and steel. Most of the larger industrials tie into the electric-distribution grid and can obtain permission to build a power plant only on the condition that the generating equipment selected has a size and heat cycle that will permit it to deliver power to the grid economically.

For example, this situation dictated a recent power installation at a steel mill of the following design conditions: A 463,000-lb-per-hr boiler, operating at 1635 psig, 977 F with reheat to 997 F. This would be an unusual cycle for most United States industrials. Similarly a chemical works last year ordered two boilers at 440,000 lb per hr, 2300 psig, and 1130 F.

These industrial units, with their dual role of generating electricity for both grid and plant use, have been built to meet high standards of service. They have permitted Sulzer Brothers to establish a good background of experience upon which United States design can make adjustments and refinements to meet American practice as it unfolds.

Author's Closure

The comments of Messrs. Allen and Powell are appreciated To the author's knowledge, there are at least five once-through units now in design or construction. Some of these are in the subcritical pressure range. To meet these different cycle conditions, a wide variety of equipment configurations have been selected.

This is a healthy development, since it will provide a broad base of operating experience with once-through steam generators within a relatively short time. Such experience will, of course, be invaluable to the entire industry in designing future units in the pressure range of 2400–5000 psi.

While it is too early to detect any conclusive trend at this time, it may be noteworthy that those companies which selected 4500–5000 psi initial pressures for their first supercritical units have already chosen 3500 psi for subsequent units now on the boards.

Advances in techniques for demineralization of water have made possible the high water purities required for supercritical pressure steam generation. To a considerable extent, these advances were brought about through the impetus of the atomic energy program. Many of their rigid requirements are being incorporated into the design of feedwater treatment equipment and materials.

It seems safe to predict that in the future there will be extensive discussion of the many problems now being encountered in the design of once-through and supercritical installations. It is only through such exchange of engineering and operating results that the maximum benefits of technological advancement can be achieved.

Effect of Internal Pressure on Flexibility and Stress-Intensification Factors of Curved Pipe or Welding Elbows

By E. C. RODABAUGH[1] and H. H. GEORGE,[2] LOUISVILLE, KY.

The flexibility and stress-intensification factors presently applied in piping-flexibility analysis to account for the behavior of curved pipe in bending have been derived from theories and tests with no internal pressure. Pressure tends to reduce the effect of these factors but in smaller and relatively thick-wall piping commonly used in the past the effect is of a low order and may be neglected; in larger diameter relatively thin-wall piping the effect is pronounced and significant. Using strain-energy methods the present paper develops a theory establishing the flexibility and stresses due to in-plane and out-of-plane bending including the effect of internal pressure, and proves its adequacy by means of carefully conducted tests. In a final step, the complex theoretical formulas are reduced to a simple and readily usable approximation.

IT IS generally known that curved pipe subjected to bending is more flexible and has higher stresses than would be indicated by the elementary theory of bending. These characteristics of curved pipe are recognized in piping-system flexibility calculations by the use of "flexibility factors" and "stress-intensification factors" which are simply the ratios of actual flexibility and stress to those predicted by the elementary bending theory.

Under the present rules of the American Standard Code for Pressure Piping (ASA B31.1-1955, Par. 621d) the flexibility factor and stress-intensification factor are given by the formulas

$$k = \frac{1.65}{h}, \geqq 1 \quad \ldots\ldots [1]$$

$$i = \frac{0.9}{h^{2/3}}, \geqq 1 \quad \ldots\ldots [2]$$

where

k = flexibility factor
i = stress-intensification factor
h = flexibility characteristic $= \frac{tR}{r^2} = \left(\frac{t}{r}\right)\left(\frac{R}{r}\right)$
t = pipe-wall thickness, in.
R = bend radius of curved pipe, in.
r = mean cross-section radius of curved pipe, in.

Equations [1] and [2] are based on theories and tests without internal pressure; however, consideration has been given in the past to the effect of internal pressure on these factors but it was concluded that the effect was negligible. For relatively thick-wall pipe used at rather low stresses, this is true; however, the recent trend has been toward the increasing use of thin-wall pipe at high stresses and in such piping the effect of the internal pressure becomes significant.

In the present paper, theoretical formulas including the effect of internal pressure are developed for flexibility and stress-intensification factors; the development is given in Appendix 1. In addition, considerable experimental work has been done to substantiate the theoretical developments and to demonstrate conclusively that internal pressure has a pronounced effect on the flexibility and stress-intensification factors; this is described in Appendix 2.

The theoretical formulas developed in Appendix 1 are extremely complex and not easily used by the practical piping engineer. It was found, however, that these complex theoretical formulas could be adequately approximated by the following readily usable formulas

$$k_p = \frac{k}{1 + \frac{S}{E} X_k}, \geqq 1 \quad \ldots\ldots [3]$$

$$i_p = \frac{i}{1 + \frac{S}{E} X_i}, \geqq 1 \quad \ldots\ldots [4]$$

where

k_p = flexibility factor with internal pressure
i_p = stress-intensification factor with internal pressure (*Note:* This factor does not include the stress caused directly by internal pressure; it covers that due to bending only.)
$S = Pr/t$ = stress due to internal pressure in straight pipe, psi
P = internal pressure, psi
E = modulus of elasticity of pipe material, psi
X_k = function of r/t and $R/r = 6\left(\frac{r}{t}\right)^{4/3}\left(\frac{R}{r}\right)^{1/3}$
X_i = function of r/t and $R/r = 3.25\left(\frac{r}{t}\right)^{2/3}\left(\frac{R}{r}\right)^{2/3}$

The effect of internal pressure on reducing the flexibility and stress-intensification factors on a curved pipe is illustrated in Figs. 1 and 2, respectively. The example covers long-radius welding elbows ($R/r = 3$) of r/t-ratios ranging from 3 to 50 and the range of internal pressures which produce up to 40,000 psi stress in steel pipe. When h is small the effect of pressure is significant. For example, without internal pressure a 24-in. standard-weight long-radius elbow ($h = 0.094$) has a flexibility factor of 17.5; with in-

[1] Senior Research Engineer, Tube Turns, a Division of National Cylinder Gas Company. Assoc. Mem. ASME.
[2] Manager of Research Section, Tube Turns, a Division of National Cylinder Gas Company.

Contributed by the Power Division and presented at the Semi-Annual Meeting, Cleveland, Ohio, June 17–21, 1956, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Brief of manuscript received at ASME Headquarters, April 6, 1955; final manuscript received, April 2, 1956. Paper No. 56—SA-50.

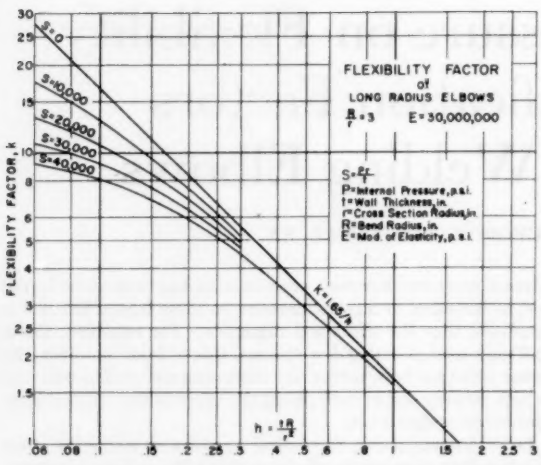


FIG. 1 EFFECT OF PRESSURE ON FLEXIBILITY FACTORS

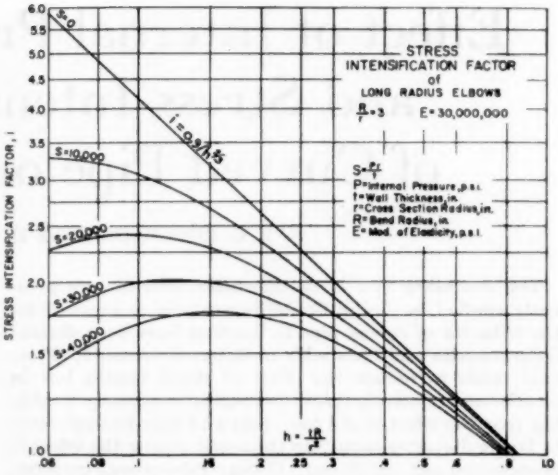
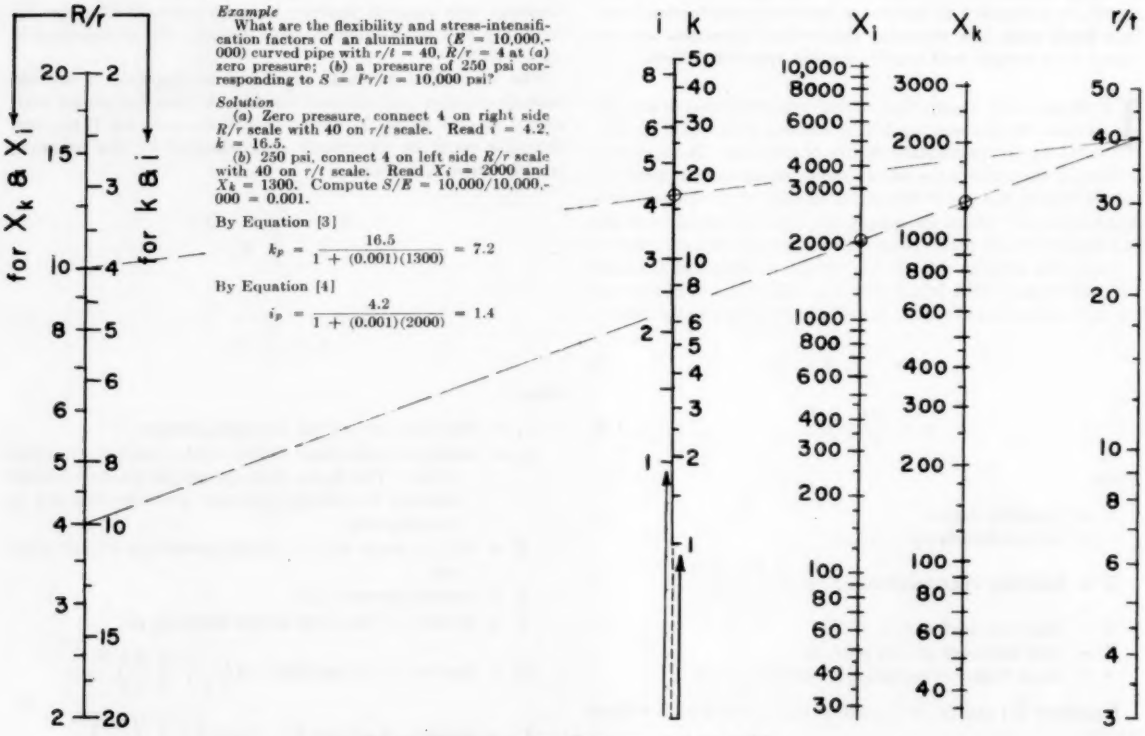


FIG. 2 EFFECT OF PRESSURE ON STRESS-INTENSIFICATION FACTORS



Example

What are the flexibility and stress-intensification factors of an aluminum (E = 10,000,000) curved pipe with r/t = 40, R/r = 4 at (*a*) zero pressure; (*b*) a pressure of 250 psi corresponding to $S = Pr/t$ = 10,000 psi?

Solution

(*a*) Zero pressure, connect 4 on right side R/r scale with 40 on r/t scale. Read i = 4.2, k = 16.5.

(*b*) 250 psi, connect 4 on left side R/r scale with 40 on r/t scale. Read X_i = 2000 and X_k = 1300. Compute S/E = 10,000/10,000,000 = 0.001.

By Equation [3]

$$k_p = \frac{16.5}{1 + (0.001)(1300)} = 7.2$$

By Equation [4]

$$i_p = \frac{4.2}{1 + (0.001)(2000)} = 1.4$$

FIG. 3 NOMOGRAPH FOR OBTAINING k, i, k_p, *and* i_p

ternal pressure corresponding to a stress of 10,000 psi the flexibility factor drops to 13.3, and at 40,000 stress to 8.2. The stress-intensification factor drops from 4.3 (S = 0) to 3.1 (S = 10,000) and to 1.7 (S = 40,000).

It is quite understandable that the effect of the internal pressure on thick-wall curved pipe was overlooked because, for values of the characteristic h covered by thick-wall pipe the change in flexibility and stress intensification with usual pressures is of the same order as might occur due to commercial variation in pipe-wall thickness.

To provide the engineer with a means for readily calculating the factors, a simple nomograph is given in Fig. 3 for obtaining k, i, X_k, and X_i as defined in Equations [1], [2], [3], and [4], respectively.

Acknowledgment

The authors wish to acknowledge their indebtedness to Mr. A. R. C. Markl for encouragement and suggestions.

BIBLIOGRAPHY

1 "Über die Formänderung dünnwandiger Rohre, insbesondere federnder Ausgleichrohre," by Th. von Karman, *Zeitschrift des Vereines deutscher Ingenieure*, vol. 55, 1911, pp. 1889–1895.

2 "Elastic Properties of Curved Tubes," by Irwin Vigness, Trans. ASME, vol. 65, 1943, pp. 105–120.

3 "Étude de la déformation et des tensions internes des Tuyaux a Ligne Moyenne Plane, Soumis a des efforts extérieurs et a une pression interne," by J. Barthélemy, Bulletin de l'Association Technique Maritime, 1947.

4 "Stresses and Deformations of Toroidal Shells of Elliptical Cross Section," by R. A. Clark, T. I. Gilroy, and E. Reissner, *Journal of Applied Mechanics*, Trans. ASME, vol. 74, 1952, pp. 37–48.

5 "Stiffness of Curved Circular Tubes With Internal Pressure," by P. G. Kafka and M. B. Dunn, *Journal of Applied Mechanics*, Trans. ASME, vol. 78, 1956, pp. 247–254.

6 "Experiments on Short-Radius Pipe Bends," by Nicol Gross, Proceedings of The Institution of Mechanical Engineers, vol. 1, 1952–1953, series B, pp. 465–479.

7 "Properties of Thin-Walled Curved Tubes of Short-Bend Radius," by T. E. Pardue and Irwin Vigness, Trans. ASME, vol. 73, 1951, pp. 77–87.

8 "Bending of Curved Thin Tubes," by Leon Beskin, *Journal of Applied Mechanics*, Trans. ASME, vol. 67, 1945, pp. A-1–7.

9 "Bending of Curved Tubes," by R. A. Clark and E. Reissner, Advances in Applied Mechanics, Academic Press, New York, N. Y., vol. 2, 1951, pp. 93–133.

10 "Fatigue Tests of Piping Components," by A. R. C. Markl, Trans. ASME, vol. 74, 1952, pp. 287–303.

Appendix 1

THEORETICAL DEVELOPMENT

Nomenclature

The following nomenclature is used in the paper:

- r = mean cross-section radius of curved pipe, in.
- t = pipe-wall thickness, in.
- R = bend radius of curved pipe, in.
- I = moment of inertia of pipe cross section = $\pi r^3 t$, in.^{4}
- E = modulus of elasticity, psi
- ν = Poisson's ratio
- M = applied moment, in-lb (M_i, in-plane; M_o, out-of-plane)
- U = energy per unit length of center line of curved pipe, lb
- P = internal pressure, psi (subscript p denotes value of a factor with internal pressure)
- S = circumferential stress in straight pipe due to internal pressure = Pr/t, psi
- α = pipe bend arc, radians
- ϕ = circumferential location angle
- w = displacement of pipe wall, in. (w_r, radial; w_t, tangential)
- ρ = radius of curvature, in a plane perpendicular to a plane containing bend radius, caused by an out-of-plane bending moment
- η = $\Delta\alpha/\alpha$ for in-plane bending, $\eta = R/\rho$ for out-of-plane bending
- ΔA = increase in pipe cross-sectional area, sq in.
- λ = $tR/r^2\sqrt{(1-\nu^2)} = h/\sqrt{(1-\nu^2)}$, flexibility characteristic
- ψ = $PR^2/Ert = SR^2/Er^2$, parameter related to pressure
- k = flexibility factor
- i = stress-intensification factor based on bending-fatigue tests
- e = strain (e_l, longitudinal; e_c, circumferential)
- S_m = stress due to applied moment, psi (S_{ml}, longitudinal; S_{mc}, circumferential)

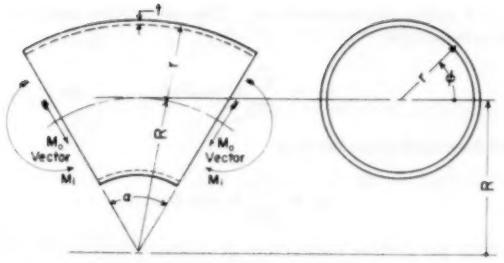


Fig. 4 Illustrating the Nomenclature

- γ = maximum circumferential stress-intensification factor (γ_i, in-plane; γ_o, out-of-plane)
- β = maximum longitudinal stress-intensification factor (β_i, in-plane; β_o, out-of-plane)

Introduction

Theoretical studies of the flexibility and stress-intensification factors of curved pipe began some 45 years ago with the work of Th. von Karman (1),[3] who developed theoretical formulas for in-plane bending of curved pipe without internal pressure. Vigness (2) developed a similar theory for the case of out-of-plane bending. A large amount of experimental work has been done which confirms their theoretical developments.

In 1947 Barthélemy (3) developed the theory for in-plane bending with internal pressure. This paper, however, apparently escaped attention in this country. Clark, Gilroy, and Reissner (4) developed a theory from which the pressure effect could be deduced. Kafka and Dunn (5) have directly developed the theory and have run confirming experiments.

None of the existing developments covers the case of out-of-plane bending nor has one developed simplified formulas which make the theory readily usable by piping engineers. In the following, the theory is developed for both in-plane and out-of-plane bending. It is then shown that the complex theoretical formulas can be represented by comparatively simple formulas with adequate accuracy for purposes of practical piping-stress analysis.

Basic Equations and Development

The effect of internal pressure can be obtained by a relatively simple extension of the energy methods used by von Karman and Vigness. Accordingly, we start with the basic energy equations developed by these authors.

In-Plane Bending

$$U_i = \frac{rtE}{2R^2}\left\{\int_0^{2\pi}\left(\frac{r\Delta\alpha}{\alpha}\sin\phi + w_t\cos\phi + w_r\sin\phi\right)^2 d\phi + \frac{t^2R^2}{12r^4(1-\nu^2)}\int_0^{2\pi}\left(\frac{d^2w_t}{d\phi^2} + \frac{dw_t}{d\phi}\right)^2 d\phi\right\} \quad [5]$$

Out-of-Plane Bending

$$U_i = \frac{rtE}{2R^2}\left\{\int_0^{2\pi}\left(\frac{rR}{\rho}\cos\phi + w_t\cos\phi + w_r\sin\phi\right)^2 d\phi + \frac{t^2R^2}{12r^4(1-\nu^2)}\int_0^{2\pi}\left(\frac{d^2w_t}{d\phi^2} + \frac{dw_t}{d\phi}\right)^2 d\phi\right\} \quad [6]$$

Equations [5] and [6] give the elastic energy stored in a unit center-line length of curved pipe due to tangential displacements

[3] Numbers in parentheses refer to the Bibliography at the end of the paper.

w_r and radial displacements w_t. The assumption made for in-plane bending is

$$w_t = \sum_{n=1}^{\infty} a_n \sin 2n\phi \quad \ldots\ldots [7]$$

and for out-of-plane bending

$$w_t = \sum_{n=1}^{\infty} b_n \cos 2n\phi \quad \ldots\ldots [8]$$

With the further assumption of inextensibility in the transverse direction, which implies that $w_r = -dw_t/d\phi$, and substituting the trigonometric series expressions for w_t, Equations [5] and [6] become

In-Plane Bending

$$U_1 = \frac{rtE}{2R^2}\left\{\int_0^{2\pi}\left(\frac{r\Delta\alpha}{\alpha}\sin\phi + \cos\phi\sum_{n=1}^{\infty} a_n \sin 2n\phi - \sin\phi\sum_{n=1}^{\infty} 2na_n\cos 2n\phi\right)^2 d\phi + \frac{\lambda^2}{12}\int_0^{2\pi}\left(-\sum_{n=1}^{\infty} 8n^3 a_n \cos 2n\phi + \sum_{n=1}^{\infty} 2na_n\cos 2n\phi\right)^2 d\phi\right\} \ldots [9]$$

Out-of-Plane Bending

$$U_1 = \frac{rtE}{2R^2}\left\{\int_0^{2\pi}\left(\frac{rR}{\rho}\cos\phi + \cos\phi\sum_{n=1}^{\infty} b_n\cos 2n\phi + \sin\phi\sum_{n=1}^{\infty} 2nb_n\sin 2n\phi\right)^2 d\phi + \frac{\lambda^2}{12}\int_0^{2\pi}\left(\sum_{n=1}^{\infty}(2n)^2 b_n\sin 2n\phi - \sum_{n=1}^{\infty} 2nb_n\sin 2n\phi\right)^2 d\phi\right\} \ldots [10]$$

The first integral in Equations [9] and [10] may be put in Fourier series form by use of trigonometric equivalents of the type

$$\cos\phi\sum_{n=1}^{\infty} a_n\sin 2n\phi = \frac{a_1}{2}\sin\phi + \frac{1}{2}\sum_{n=1}^{\infty}(a_n + a_{n+1})\sin(2n+1)\phi$$

Performing the indicated integrations in Equations [9] and [10], it follows that both Equations [9] and [10] become

$$U_1 = \frac{\pi rtE}{2R^2}\left\{r^2\eta^2 + 3r\eta c_1 + \frac{9}{4}c_1^2 + \frac{1}{4}\sum_{n=1}^{\infty} c_n^2(1-2n)^2 - 2c_nc_{n+1}(2n-1)(2n+3) + c_{n+1}^2(2n+3)^2 + \frac{\lambda^2}{12}\sum_{n=1}^{\infty} c_n^2(8n^3-2n)^2\right\} \ldots [11]$$

where $c_n = a_n = b_n$; $\eta = \Delta\alpha/\alpha$ for in-plane bending; $\eta = R/\rho$ for out-of-plane bending.

When internal pressure is considered, the additional work represented by the internal pressure P acting against the change in volume must be considered, per unit length of center line, this is

$$U_2 = P\Delta A \quad \ldots\ldots [12]$$

ΔA, the increase in area of the curved pipe cross section is

$$\Delta A = \frac{1}{2}\int_0^{2\pi}(r + w_r)^2 d\phi - \frac{1}{2}\int_0^{2\pi} r^2 d\phi = \frac{1}{2}\int_0^{2\pi}(2rw_r + w_r^2)d\phi \ldots [13]$$

The relation $w_r = -dw_t/d\phi$ is derived from the condition of inextensibility, dropping second and higher-order terms. In the case of determining ΔA, however, this is not sufficiently accurate, since the work done by the pressure is itself dependent upon second and higher-order terms in ΔA. Considering second-order terms, it can be shown that

$$\Delta A = -2\pi\sum_{n=1}^{\infty} n^2(4n^2-1)c_n^2 \quad \ldots\ldots [14]$$

In order to determine the values of the coefficients c_n, we may differentiate $U = U_1 - U_2$ with respect to each c_n and, by the principle of least work, each of the resulting expressions may be set equal to zero, thereby obtaining

$$\left.\begin{aligned}
\frac{\partial U}{\partial c_1} &= 0 = 3r\eta + (5 + 6\lambda^2 + 24\psi)c_1 - \frac{5}{2}c_2 \ldots\ldots \\
\frac{\partial U}{\partial c_2} &= 0 = -\frac{5}{2}c_1 + (17 + 600\lambda^2 + 480\psi)c_2 - \frac{21}{2}c_3 \ldots\ldots \\
\frac{\partial U}{\partial c_n} &= 0 = -c_{n-1}\frac{(2n-3)(2n+1)}{2} + c_n\left\{(4n^2+1) + (8n^3-2n)^2\frac{\lambda^2}{6} + [8n^2(4n^2-1)]\psi\right\} - c_{n+1}\frac{(2n-1)(2n+3)}{2} \ldots.
\end{aligned}\right\} \ldots [15]$$

Equations [15] give a set of n linear equations with $(n + 1)$ unknown c's. By assuming that $c_{n+1} = 0$, all constants may be evaluated. Since the c's all contain a factor of $r\eta$, it is convenient to introduce the relation $d_n = c_n/r\eta$.

The minimized energy U is then equated to work done by the bending moment

$$U_{\min} = \frac{1}{2}\frac{M\eta}{R} \quad \ldots\ldots [16]$$

Flexibility Factor

Substituting the values of $d_n = c_n/r\eta$ from Equations [15] in Equation [16] and solving for η

$$\eta = \frac{RM}{IE}\left[1 + 3d_1 + \frac{9}{4}d_1^2 + \frac{1}{4}\sum_{n=1}^{\infty} d_n^2(1-2n)^2 - 2d_nd_{n+1}(2n-1)(2n+3) + d_{n+1}^2(2n+3)^2 + \frac{\lambda^2}{12}\sum_{n=1}^{\infty} d_n^2(8n^3-2n)^2 + \psi\sum_{n=1}^{\infty} 4n^2(4n^2-1)d_n^2\right]^{-1} \ldots [17]$$

The factor in brackets in Equation [17] is the generalized formula for the flexibility factor k for both in-plane and out-of-plane bending. It may be simplified to the form

$$k_p = \frac{1}{1 + \frac{3}{2} d_1} \quad \ldots\ldots [18]$$

In the first approximation, for example, with c_2 and higher c_n assumed zero, Equation [15] gives

$$d_1 = \frac{c}{r\eta} = -\frac{3}{5 + 6\lambda^2 + 24\psi} \quad \ldots\ldots [19]$$

and from Equation [18]

$$k_p = \frac{1}{1 - \frac{3}{2}\left(\frac{3}{5 + 6\lambda^2 + 24\psi}\right)} = \frac{5 + 6\lambda^2 + 24\psi}{0.5 + 6\lambda^2 + 24\psi} \quad \ldots [20]$$

When $P = 0$, Equation [20] reduces to the equations derived by von Karman (in-plane bending) and Vigness (out-of-plane bending). Equation [20], when $P \neq 0$, is confirmed by the work of Kafka and Dunn for in-plane bending by substituting $\nu = 0$ in their equations. Note, however, that the development herein includes the flexibility factor with internal pressure for out-of-plane bending. It turns out that, as in the case with zero pressure, the flexibility factor for in-plane bending is the same as for out-of-plane bending.

Stresses

The longitudinal strains on which Equations [5] and [6] are based are as follows:

In-Plane Bending

$$e_l = \frac{1}{R}\left(\frac{\Delta\alpha}{\alpha} r \sin\phi + w_l \cos\phi + w_r \sin\phi\right) \quad \ldots\ldots [21]$$

Out-of-Plane Bending

$$e_l = \frac{1}{R}\left(\frac{Rr}{\rho}\cos\phi + w_l \cos\phi + w_r \sin\phi\right) \quad \ldots\ldots [22]$$

Circumferential strains at the outer and inner-wall surfaces, for both in-plane and out-of-plane bending, are given by

$$e_c = \pm \frac{t}{2(1 - \nu^2)r^2}\left(\frac{d^2 w_l}{d\phi^2} + \frac{dw_l}{d\phi}\right) \quad \ldots\ldots [23]$$

The plus sign applies to the outside-wall surface; minus to inside-wall surface.

Strains may be converted to stresses by the usual formulas

$$S_{ml} = \frac{E}{1 - \nu^2}(e_l + \nu e_c) \quad \ldots\ldots [24]$$

$$S_{mc} = \frac{E}{1 - \nu^2}(e_c + \nu e_l) \quad \ldots\ldots [25]$$

Combining Equations [21] through [25] along with Equations [7] and [8] and noting that $\eta = RMk/IE$, the following equations for the stresses are obtained

In-Plane Bending

$$S_{ml} = \frac{k_p M_i r}{I(1 - \nu^2)}\left\{\left(1 + \frac{3d_1}{2}\right)\sin\phi + \frac{1}{2}\sum_{n=1}^{\infty}\{d_n(1 - 2n) + d_{n+1}(2n + 3)\}\sin(2n + 1)\phi \pm \frac{\nu\lambda}{2}\sum_{n=1}^{\infty} d_n(2n - 8n^3)\cos 2n\phi\right\} \quad \ldots [26]$$

$$S_{mc} = \frac{k_p M_i r}{I(1 - \nu^2)}\left\{\pm \frac{\lambda}{2}\sum_{n=1}^{\infty} d_n(2n - 8n^3)\cos 2n\phi + \nu\left(1 + \frac{3d_1}{2}\right)\sin\phi + \frac{\nu}{2}\sum_{n=1}^{\infty}\{d_n(1 - 2n) + d_{n+1}(2n + 3)\}\sin(2n + 1)\phi\right\} \quad \ldots [27]$$

Out-of-Plane Bending

$$S_{ml} = \frac{k_p M_o r}{I(1 - \nu^2)}\left\{\left(1 + \frac{3d_1}{2}\right)\cos\phi + \frac{1}{2}\sum_{n=1}^{\infty}\{d_n(1 - 2n) + d_{n+1}(2n + 3)\}\cos(2n + 1)\phi \pm \frac{\nu\lambda}{2}\sum_{n=1}^{\infty} d_n(8n^3 - 2n)\sin 2n\phi\right\} \quad \ldots [28]$$

$$S_{mc} = \frac{k_p M_o r}{I(1 - \nu^2)}\left\{\pm \frac{\lambda}{2}\sum_{n=1}^{\infty} d_n(8n^3 - 2n)\sin 2n\phi + \nu\left(1 + \frac{3d_1}{2}\right)\cos\phi + \frac{\nu}{2}\sum_{n=1}^{\infty}\{d_n(1 - 2n) + d_{n+1}(2n + 3)\}\cos(2n + 1)\phi\right\} \quad \ldots [29]$$

Stress-intensification factors are obtained by dividing Equations [26] through [29] by Mr/I, since these factors are the ratios of the curved-pipe stress to the calculated stress by the ordinary beam theory; which stress is simply Mr/I.

For example, the circumferential stress-intensification factor for in-plane bending by a first approximation, from Equation [27] is

$$\frac{s_{mc}}{Mr/I} = \frac{k_p}{1 - \nu^2}\left\{\mp 3d_1\lambda\cos 2\phi + \nu\left[\left(1 + \frac{3d_1}{2}\right)\sin\phi - \frac{d_1}{2}\sin 3\phi\right]\right\} \quad \ldots [30]$$

When $P = 0$, this equation reduces to that obtained by Gross (6),[4] based on von Karman's development.

The longitudinal stress-intensification factor for in-plane bending by a first approximation, from Equation [26], is

$$\frac{S_{ml}}{Mr/I} = \frac{k_p}{1 - \nu^2}\left\{\mp 3d_1\nu\lambda\cos 2\phi + \left(1 + \frac{3d_1}{2}\right)\sin\phi - \frac{d_1}{2}\sin 3\phi\right\} \quad \ldots [31]$$

When $P = 0$ this equation reduces to that obtained by Gross.[4] When $P \neq 0$, $\nu = 0$ and, considering the mid-wall stress only, this equation gives

$$\frac{S_{ml}}{Mr/I} = k_p\left\{\left(1 + \frac{3d_1}{2}\right)\sin\phi - \frac{d_1}{2}\sin 3\phi\right\} \quad \ldots\ldots [32]$$

Equation [32] is identical to that obtained by Kafka and Dunn with $\nu = 0$.

[4] Except for the factor $(1 - \nu^2)$ which is omitted from von Karman's development.

In this connection it might be noted that a number of previous papers on this subject give longitudinal stresses at the mid-wall of the pipe only. The longitudinal stresses at the inner and outer-wall surfaces are affected by the circumferential bending stresses as indicated by Fig. 8.

Assumptions of Theory

The theory is that of thin shells, wherein the wall thickness is assumed small compared to the radius of curvature. This assumption is generally true where the flexibility and stress factors are of importance. The theory assumes that the ratio R/r is large compared to unity. This is not true of welding elbows; however it has been shown, both theoretically (7) and experimentally, that even for R/r as small as 2, the theory is still of adequate accuracy.

It is assumed that the curved pipe or elbow cross section is initially round. Slight deviations from roundness have no significant effect on the theory as given herein; however, it might be noted that if a curved pipe or elbow is slightly out of round, application of pressure will produce rotation of the elbow ends (if an elbow end is free) or end moments (if the elbow ends are restrained).

The results obtained herein theoretically apply to a curved pipe of any bend arc, α. Actually, the end effects of the attached pipe or flanges can have a pronounced effect for small bend arcs combined with small values of R/r, as shown experimentally by Pardue and Vigness (7).

Convergence

In series-type equations such as those given herein, it is necessary to know how many terms in the series must be used in order to obtain satisfactory accuracy. The rapidity of convergence of the series expressions for the various factors is not uniform; convergence is fastest for the longitudinal stresses, followed by the flexibility factor and slowest for the circumferential stresses. Since the circumferential stress is usually higher than the longitudinal stress, the number of terms required to obtain satisfactory accuracy of the circumferential stress governs the degree of approximation required.

In the case of zero pressure, the flexibility and stress-intensification factors depend only upon λ. Comparison of values by successive approximations indicates that the order of approximation given in Table 1 is necessary for accuracy within about 10 per cent.

TABLE 1 ORDER OF APPROXIMATION

Value of λ	Approximation order required
0.5 and larger	1st
0.4 to 0.16	2nd
0.12 to 0.08	3rd
0.06 to 0.04	4th

With internal pressure, the series converges somewhat faster than indicated above; however, in general, for thin-wall, short-bend-radius elbows a relatively high order of approximation is necessary for satisfactory accuracy of the circumferential stress-intensification factor.

Application of the series formulas to long-radius welding elbows ($R/r = 3$) is shown in Fig. 5. In calculations for Fig. 5, an attempt was made to obtain accuracy within about ± 2 per cent. It was necessary to use a fifth-order approximation for $\lambda = 0.03$ to attain this accuracy.

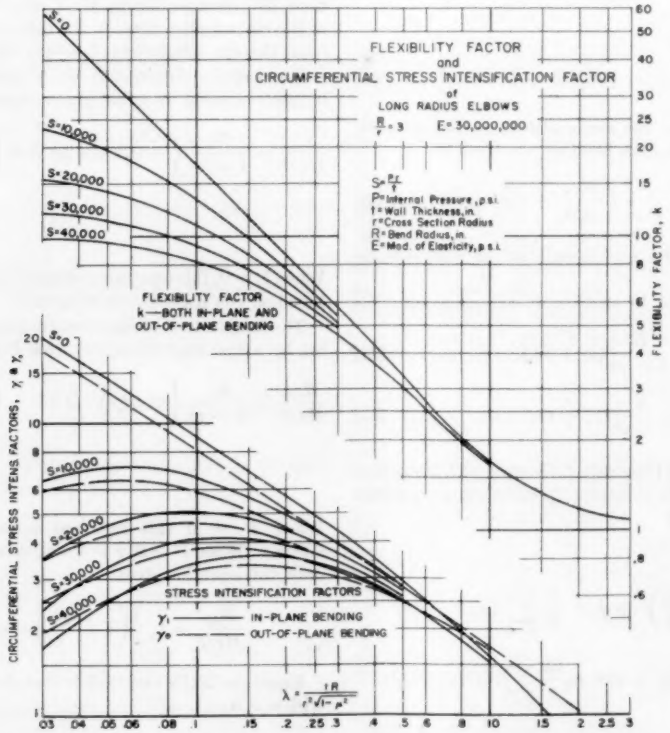


FIG. 5 SERIES SOLUTIONS FOR FLEXIBILITY AND STRESS-INTENSIFICATION FACTORS, LONG-RADIUS WELDING ELBOWS

Approximate Formulas

It will be apparent, in writing out Equations [18], or [26] through [29], that expressions for the third or fourth-order approximation become quite lengthy and time-consuming to apply. In the case of zero pressure, it was found graphically by Beskin (8) and later shown analytically by Clark and Reissner (9) that, for values of λ less than about 0.3, the flexibility and maximum stress-intensification factors can be expressed by the following simple formulas

$$k = \frac{1.73}{\lambda} \quad [33]$$

$$\gamma_i = \frac{1.95}{\lambda^{2/3}} \quad [34]$$

$$\gamma_o = \frac{1.67}{\lambda^{2/3}} \quad [35]$$

where k, γ_i, γ_o, are the flexibility factor, in-plane bending maximum circumferential stress-intensification factor, and out-of-plane bending maximum circumferential stress-intensification factor, respectively; all at zero pressure.[5]

The ASA B31.1 Code formula for k (Equation [1] herein) is simply Equation [33] using the parameter $h = tR/r^2$ instead of λ, i.e.

$$k = \frac{1.73}{\lambda} = \frac{1.73\sqrt{(1-\nu^2)}}{h} = \frac{1.65}{h}$$

where ν (Poisson's ratio) is taken as 0.3.

The ASA B31.1 Code formula for i (Equation [2] herein) is approximately one half of Equation [34]. This is based on extensive bending-fatigue tests of elbows (10). It was found that using the fatigue life of butt-welded straight pipe as unity, the effective stress-intensification factors of elbows in bending fatigue were about one half of the theoretical value given by Equations [34] and [35]. The fatigue tests showed that the stress-intensification factor for in-plane bending was slightly higher than for out-of-plane bending, as indicated by a comparison of Equation [34] with Equation [35]; however, this difference is rather small for practical significance. Therefore, in the interest of simplicity, the higher i-value is given in the code for application to both in-plane and out-of-plane bending.

The existence of simple asymptotic formulas for the flexibility and maximum-stress factors without pressure suggested the possibility of similar formulas for these factors with pressure. Values of k_p, γ_{ip}, and γ_{op} were calculated by the series formulas over an appropriate range of the variables S/E, r/t, and R/r. In the case of γ_{ip}, the maximum stress occurs at $\phi = 0$. In the case of γ_{op}, the maximum stress lies between 0 and 45 deg; its value and location were determined by plotting the results of calculations for various values of ϕ in this region. As is the case for zero pressure, the value of γ_{op} was found generally to be less than γ_{ip}. Accordingly, approximate formulas were developed for k_p and γ_{ip} only.

It was found that by plotting the following

$$\frac{k}{k_p} - 1 \quad \text{and} \quad \frac{\gamma_i}{\gamma_{ip}} - 1$$

against the variables S/E, r/t, and R/r on logarithmic co-ordinates, that substantially straight-line relations held over the range of these variables likely to be encountered in piping practice; i.e., S/E corresponding to stresses up to 40,000 psi in steel pipe; r/t up to 50 (24 in. Schedule-10 pipe) and R/r from 2 to ∞. These plots then led to the approximate formulas

$$k_p = \frac{k}{1 + 6\dfrac{S}{E}\left(\dfrac{r}{t}\right)^{4/3}\left(\dfrac{R}{r}\right)^{1/3}} = \frac{k}{1 + \dfrac{S}{E}X_k} \quad [36]$$

$$\gamma_{ip} = \frac{\gamma_i}{1 + 3.25\dfrac{S}{E}\left(\dfrac{r}{t}\right)^{2/3}\left(\dfrac{R}{r}\right)^{2/3}} = \frac{\gamma_i}{1 + \dfrac{S}{E}X_i} \quad [37]$$

[5] The maximum longitudinal stress-intensification factor is of less practical interest since it generally is smaller than the circumferential factor. For small values of λ, $P = 0$, the ratio of the maximum circumferential stress to the maximum longitudinal stress is about 2 for in-plane bending and about $1^1/_2$ for out-of-plane bending.

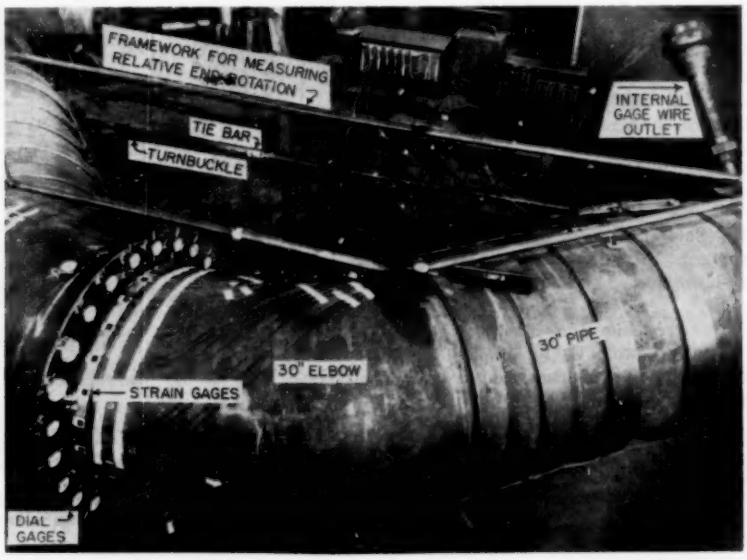


Fig. 6 Testing Arrangement, 30-In. Long-Radius Welding Elbow

These formulas[6] have been found to be within 10 per cent of the theoretical results over the range of variables included where the Equations [33], [34], and [35] are themselves accurate. In the regions where the approximate formulas are not in good agreement with theory, the flexibility and stress-intensification factors are low and of lesser significance in practical piping applications.

Appendix 2

COMPARISON OF THEORY WITH TEST DATA

Like many other developments in the field of theoretical applied mechanics, the theory presented herein is based on a number of simplifying assumptions which enable a relatively simple solution to be obtained. Accordingly, it is desirable to check the theory with tests to make sure that the theoretical development is of adequate accuracy. Numerous investigators have checked theory against test at zero pressure and found adequate agreement. The tests described herein, therefore, are principally concerned with comparing the theoretical pressure effect with experimental results.

Tests were run on a 30-in-OD, 0.500-in. nominal wall welding elbow with a bend radius of 45 in. The average wall thickness of the elbow was 0.515 in. with a variation of +0.058 in., —0.094 in. The average outside diameter was 29.973 in. with a variation of +0.160 in., —0.172 in. The elbow was made of a low-alloy, high-yield-strength proprietary steel.

Each end of the elbow was welded to 59-in. lengths of 30-in-OD, 0.500-in. nominal wall pipe purchased to API Standard 5LX52. The open ends of these two pipe lengths were closed with 30-in. welding caps, with a flanged opening in one cap for access to inside strain gages. An over-all view of the test assembly is shown in Fig. 6.

In-plane bending moments were applied by means of a tie rod hinged to each cap with a turnbuckle in the tie rod as shown in Fig. 6. Loads were measured by SR-4 strain gages on the tie rod. The rotation of one end of the elbow with respect to the other was measured by means of a framework attached to the straight pipe with dial gages to measure relative movement; the framework can be seen in Fig. 6. In calculating the flexibility factor from these measurements, bending of the 30-in. straight pipe was taken into consideration.

The stress in the elbow wall was determined by means of SR-4 strain gages placed around the circumference every 15 deg from crotch ($\phi = -90°$) to back ($\phi = 90°$), plus a check gage on the opposite side ($\phi = 180°$). Strain gages were placed on both the inside and outside-wall surfaces.

The tests consisted of applying loads with the turnbuckle in the tie rod in steps up to 30,000 lb, measuring deformations and strains at each step. This was repeated at internal pressures of 0, 400, 800, and 1100 psi. Several runs were made for each pressure and the results averaged. Pressure was applied with water and, subsequently, with nitrogen gas. Inside strain gages were used with nitrogen gas only.

The flexibility factors and maximum stress-intensification factors obtained in these tests are shown in the top graph of Fig. 7. The pattern of stress around the circumference and its variation with pressure are shown in Fig. 8.

[6] More accurate formulas can be established, however; in such formulas the exponent of S/E, instead of being unity, is a function of r/t. For practical piping work, the slight gain in accuracy is not believed to compensate for the resulting increase in complexity of the formulas.

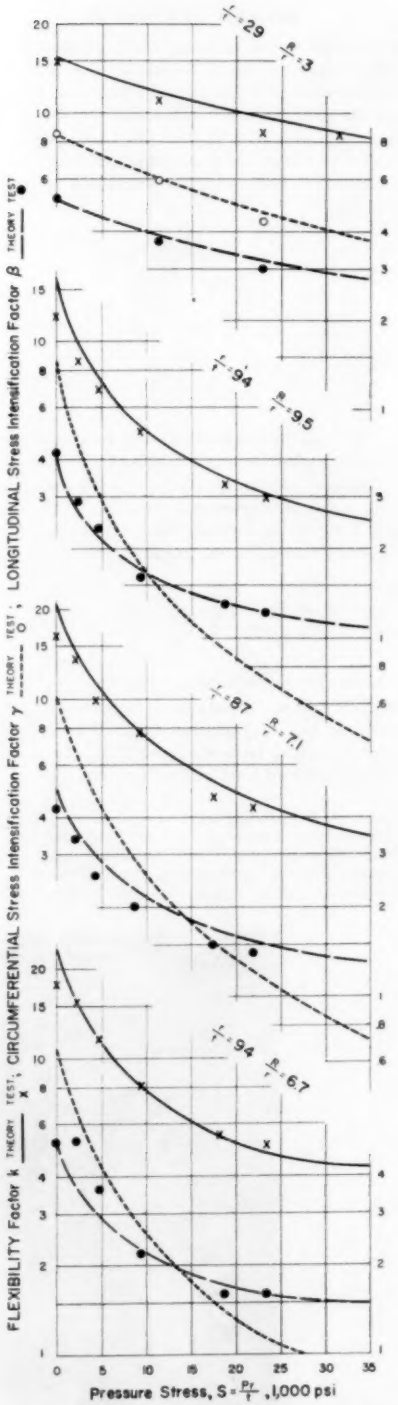


FIG. 7 CHANGE IN FLEXIBILITY FACTOR AND MAXIMUM STRESS-INTENSIFICATION FACTORS DUE TO INTERNAL PRESSURE

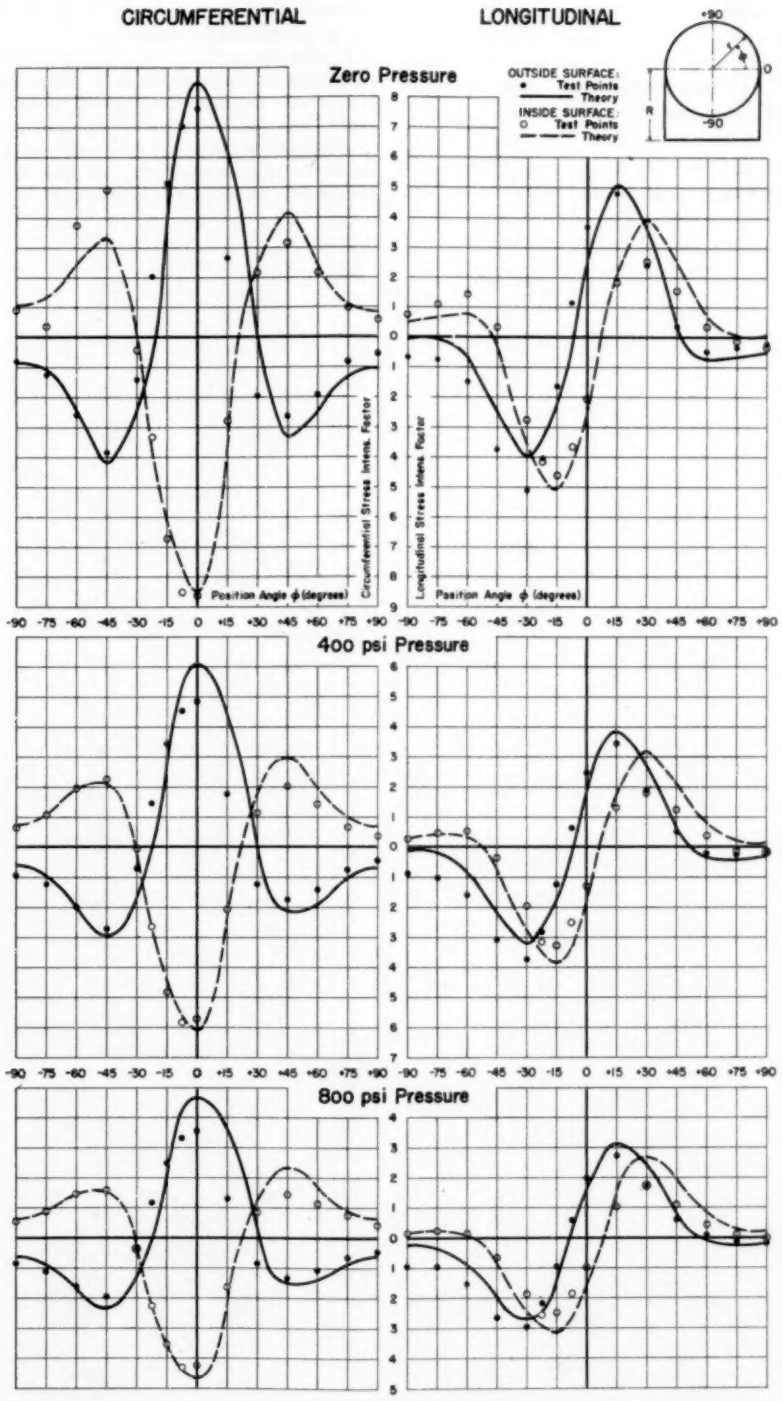


FIG. 8 VARIATION OF STRESS AROUND CIRCUMFERENCE OF AN ELBOW OF 30-IN. OD, 0.515-IN. AVERAGE WALL, 45-IN. BEND

Results of similar in-plane bending tests (5) on curved pipe with outside diameters of 3 to 3½ in., wall thicknesses of 0.016 to 0.020 in., and bend radii of 10 in. to 14¼ in. are shown in the lower three graphs of Fig. 7. Circumferential stress-intensification factors were not reported in these tests.

In general, the theory and tests are in adequate agreement. The fact that the tests and theory do not agree precisely is probably due in part to such experimental factors as nonuniform wall thickness and end effects. In addition, the theory itself is not precise as pointed out in Appendix 1.

It is interesting to note that the test results shown cover a wide range of curved-pipe dimensions. Agreement of theory and test is adequate for these dimensionally extremely different curved pipes. It is also of interest to note that the approximate Equation [33], while not intended to apply for values of r/t much over 50, is still in adequate agreement with theory and tests for the curved pipe where r/t was as high as 94.